DOCUMENT RESUME

ED 022 496 24 LI 000 815

By-Langridge, D.W.; And Others
INDEXING FOR ERIC. VOLUME 2, LESSON 1 - INTRODUCTION TO INDEXING, LESSON 2 - HOW TO INDEX A DOCUMENT.
Maryland Univ., College Park. School of Library and Information Services.
Spons Agency-Office of Education (DHEW), Washington, D.C. Bureau of Research.
Bureau No-BR-6-1837
Pub Date Aug 67
Grant-OEG-2-6-061837-0779
Note-208p.
EDRS Price MF-$1.00 HC-$8.40
Descriptors-*COORDINATE INDEXES, DOCUMENTATION, INDEXES (LOCATERS), *INDEXING, INFORMATION RETRIEVAL, INSTRUCTIONAL MATERIALS, PROGRAMED MATERIALS, PROGRAMED TEXTS, SEARCH STRATEGIES, *THESAURI, VOCABULARY
Identifiers-*Educational Resources Information Center, ERIC

This volume contains the first two lessons of a course in subject indexing based upon the Educational Resources Information Center (ERIC) Thesaurus. Lesson 1 is an introduction to indexing, containing definitions and concepts related to the theory of indexing and information retrieval as well as an introduction to the ERIC Thesaurus. Lesson 2 is about the procedures involved in indexing a document. In this lesson the problems of choosing useful indexing concepts and translating them into terms used in the ERIC Thesaurus are dealt with. Also included is a discussion of the hierarchical relationships displayed in the thesaurus, the use of identifiers, and major descriptors. The last two lessons of the course are contained in volume 3 (LI 000 816) and the appendices are in volume 1 (LI 000 815). (CM)

LI 000 815

DR6-1837
PA-24

ERIC/CLIS
JUL 23 '68
UNIVERSITY OF MINNESOTA

INDEXING FOR ERIC

Vol. 2

ED022496

LI 000 815

INDEXING FOR ERIC

VOLUME 2

Lesson 1 — INTRODUCTION TO INDEXING
Lesson 2 — HOW TO INDEX A DOCUMENT

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Produced at: The School of Library and Information Sciences, University of Maryland, under grant from U. S. Office of Education.

Principal Investigators: D. W. Langridge, J. Mills, J. Perreault

Programmed by: Computer Applications Incorporated, Silver Spring, Maryland, I. Goldberg, C. G. Davis, Mercedes Phillips, L. Frandsen

Consultant: C. D. Batty

Lesson 1
INTRODUCTION TO INDEXING

INTRODUCTION

Indexing is not a new activity: it has been practiced since libraries came into existence. Modern information services are in essence the same as traditional libraries. Specific problems and emphases may change, but the fundamental principles of indexing remain the same.

The purpose of indexing is to show which documents may be useful in answer to a particular inquiry. Some forms of inquiry are fairly clear and straight forward; for example, those which state the author, title, publisher, or number of a document. It is important that an information service has access to its collection through these characteristics, since one or the other is often the only thing an inquirer knows about a document he wants. The standard indexing form for ERIC Centers provides for these items, and there are rules which must be carefully followed in recording them. Accuracy and consistency are essential. However, it is comparatively easy to learn this part of indexing, and this course is concerned solely with the more difficult part — subject indexing.

What is the purpose of this course?

To teach a complete course in theory and techniques of indexing.	5
To teach indexing theory and techniques of subject indexing.	9

If you are reading this page you have made a mistake.

Remember, this is not an ordinary book. The page you should turn to next is always indicated at the bottom of each page.

Please turn back to page 1. Read the page carefully and answer the question. After you have chosen the best answer, turn to the page indicated by your choice.

(from page 9)

You say that by keeping a large collection of documents, the purpose of a library will be fulfilled.

This may be necessary to fulfill the purpose of a library, but the collection itself is not the fulfillment of purpose. Information locked in a vault is of no use to anyone; it needs to be brought out where it can be used. This usefulness represents the fulfillment of the purpose of the collection.

How can we make sure documents are used to the best advantage?

By setting up a system to find documents when required.	6
By describing the subjects of documents precisely.	12

4
(from page 7)

It seems to you that consulting the library policy to determine the desired depth of detail to which a document should be indexed is the first job in indexing a document.

As an indexer, you should have a good idea of policy in regard to depth of indexing before you try to index a document whithout assistance.

Please turn back to 7 and read the section about CONCEPT INDEXING again. Then look again at the question and select another answer.

(from page 1)

You say that the purpose of this course is to teach a complete course in theory and techniques of indexing.

Not quite. A complete course in theory and techniques of indexing would include all forms of indexing, techniques of maintaining card and book files, methods of indicating author, title, and publisher, etc. Rules for indexing in ERIC indicate how to include an author, title, etc, but it is beyond the scope of this course to treat the subject in such complete detail.

This course is confined to one aspect of information retrieval; namely, the question of what the document is about.

Now continue with the course by turning to 9.

(from pages 3, 9 , or 12)

You are correct. No matter how large the collection or how precisely documents are described, the library is of no use if it does not make information available to those who need it.

The main elements of an information retrieval system are outlined in the definitions which follow.

INDEXING

Indexing is that part of information retrieval concerned with specifying the content of documents.

In this course we shall limit the term indexing to mean subject indexing. Our main concern is to describe the subject content of documents in such a way that those which deal with several different aspects of a subject, or discuss several different subjects, will be equally accessible to a searcher, regardless of which subject he requests.

A technique which is related to subject indexing is abstracting, especially in the method of examining a document to analyze its contents; but we are not concerned with it in this course. You may find that the demonstrations in the last section of the course are helpful for abstracting.

Indexing for ERIC consists of two stages — — concept indexing and translation.

Continue on 7.

(from page 6)

CONCEPT INDEXING

In concept indexing the main concepts treated in a document are identified and expressed in the indexer's own words; it is only later that he may translate these into the language of the thesaurus.

When an indexer considers the content of a document, he starts by listing the concepts as he finds them. He may use any terms he likes, either the author's or his own. At this stage of his work, he may list a great many terms to be sure that all concepts that might conceivably be used for indexing the document are included. The depth of detail to which he should go in indexing will depend upon the library policy and his own judgment. There are no set rules for this decision.

What is the indexer's first job in indexing a document?

Consulting the library policy to determine the depth of detail to which it should be indexed.	4
Expressing the concepts or ideas of the document in convenient terms	10
Expressing the concepts or ideas of the document in terms of the ERIC system.	13

8
(from page 15)

No. Summarization does not attempt to list all the important subjects of a document. Summarization is a common feature of conventional indexing in which we are more concerned with a single brief description of the whole subject of the document.

Now go on to 11, and continue with the course.

(from page 1 or 5)

Good. This course has been designed to teach you indexing theory and its applications to subject indexing for the ERIC system.

The precise discussion of a special subject often requires new terms, or old terms used with limited meanings. Unfortunately, the literature of indexing is full of unnecessary new terms, terms used loosely, and terms used by different authors with different meanings. In an attempt to minimize confusion we shall begin with definitions of the major terms used in the course.

DEFINITIONS

Information Retrieval

Information retrieval is the general function or purpose of a library; indexing is one of several operations which support this purpose.

Many definitions have been given, but we shall use the following:

Information retrieval: Recovery of all documents of a specified content from a given collection of documents.

How is the purpose of a library fulfilled?

By keeping a large collection of documents.	3
By maintaining a system for recovery of information	6
By describing documents precisely.	12

(from page 7)

You say that expressing the concepts or ideas of the document in convenient terms is the first job in indexing a document.

You are correct. The first step in indexing a document is to list the ideas or concepts in either the author's or your own words.

Once the indexer has listed the concepts contained in a document, he proceeds to the second step in indexing, namely, translation.

TRANSLATION

Translation is the conversion of the indexer's words into the controlled language of the authority list.

Once the *indexer* has listed the concepts in his own words, he will need to select from the authority list the words which it has used for the same concepts. By using the same authority list the *searcher* can define his needs in the same terms that the indexer used to define the documents.

Why must document concepts be translated?

So that the indexer and searcher can represent concepts by similar terms.	15
So that the indexer will define the author's concepts precisely.	18
So that the searcher's needs will be specified precisely.	21

(from page 8 or 15)

Good.

PRE-COORDINATION

When the summarized subject comprises more than one concept, the terms are combined when the document is indexed. This is known as pre-coordination. The example, "Schools – Administration", shows that with pre-coordination the relationship of one term with another may be clearly shown. Here the subject is administration of schools, and this document would not be retrieved by a search for any other type of administration or any other aspect of schools.

Pre-coordination of terms for subjects which consist of more than one concept takes place:

during indexing.	14
during searching.	17
during both indexing and searching.	22

(from page 9)

You say that by describing documents precisely, the purpose of a library will be fulfilled.

This is one necessary step in fulfilling the purpose of a library, but the precise descriptions alone are not sufficient. We need a way of recognizing these descriptions in answer to our questions and of retrieving the necessary documents from the collection.

Now try again.

How can we make sure documents are used to the best advantage?

By collecting a large number of documents.	3
By setting up a system to find documents when required.	6

(from page 7)

It seems to you that expressing the concepts or idea of the document in terms of the ERIC system is the first job in indexing a document.

You can't translate ideas until you have identified them! Translation will have to be done before the job is finished, but it can't be done first. To identify the ideas in a document the indexer is free to use any words or expressions, his own or the author's, that he feels will adequately describe the contents of the document.

Please turn back to 7 and read the section on CONCEPT INDEXING again. Then try the question once more.

(from pages 11, 17, or 22)

You say that in pre-coordination the terms are combined during indexing.

Yes. Pre-coordination is an essential part of conventional indexing.

COORDINATE INDEXING

Coordinate indexing is an abbreviation commonly used for post-coordinate indexing. Coordinate systems are characterized by post-coordination and depth indexing.

DEPTH INDEXING

In contrast to summarization, depth indexing attempts to extract all the main concepts dealt with in the document and for this reason it is also known as exhaustive indexing or extraction. This is a relative idea, and the depth of indexing in a system may be anywhere in the range that lies between summarization and a concordance, which indexes every word in a document. Coordinate indexing is not defined by this characteristic, but in practice these systems lend themselves very readily to depth indexing. There is no precise way of defining depth indexing, and the requirements of a particular system must be learned from experience.

Which of the following is characteristic of depth indexing?

Representing the main subject by a few terms.	20
Including all important concepts treated in documents.	29

(from page 10)

You are correct. Translating concept terms into the language of the authority list enables the indexer and searcher to use the same terms. This is a necessary condition of effective retrieval of information.

Let us now proceed to the definition of the type of indexing that is usually found in a library.

CONVENTIONAL INDEXING

In most libraries conventional indexing procedures consist of summarization and pre-coordination.

SUMMARIZATION

Summarization is the representation of the total content of a document by a brief description. This is not an inevitable characteristic of conventional systems, but it is dictated by economic considerations.

For example, a 300-page book may be indexed merely as "Education — Administration." It could have chapters on Policy, Planning, Organization, College Administration, School Administration, etc. These concepts are not explicitly stated in the indexing since they are considered to be implicit in the general experession "Education — Administration."

Which of these definitions is correct?

Summarization lists all the important subjects mentioned in a document.	8
Summarization represents the whole subject of a document.	11

(from page 29)

You think that "indexing of subjects in depth" best characterizes a post-coordinate system.

No. Most coordinate indexing systems do index in depth, but this is not the essential feature.

Return to 29, read the section on POST - COORDINATION, and consider the question again.

(from page 11)

It seems to you that pre-coordination is the combining of terms during searching.

No; the prefix "pre" should have told you that the word belongs to an activity that takes place before searching. Pre-coordination of terms is an essential part of conventional indexing in which the terms are assembled to make a single description (though perhaps a complex one) of the whole subject of the document.

Now turn to 14, and continue with the course.

(from page 10)

You think that document concepts must be translated so that the indexer will define the author's concepts precisely.

Apparently you think the authority list is a type of dictionary to help the indexer clarify his definitions of the author's concepts. It does this to some degree, but defining words is not the purpose of that authority list. It is difficult for the searcher to know what terms have been used in indexing. The authority list offers a standard that both indexer and searcher can use.

Now go back to 10, read the section on TRANSLATION again. Then reconsider the question and select the correct answer.

(from page 29)

You think that indexing the document mainly by isolated terms characterizes a post-coordinate system.

This is correct. Post-coordination leaves the combining of terms until the time of searching. The indexer describes documents mainly by isolated elementary terms and the searcher combines them in his statement of the question.

INDEXING LANGUAGE

The indexing language (also known as controlled language or authority list) is a device which brings together the language of the documents and the language of the inquiry. In coordinate indexing systems it is usually known as a thesaurus. In conventional indexing systems it takes the form of a classification scheme or subject heading list. There are two major aspects of every index language: (a) the individual terms and (b) the relationships between these terms. We shall deal with these later.

How can indexers and searchers be sure of using the same terms to express concepts indexed in ERIC?

By observing the rules for combining terms.	23
By referring to the same indexing language.	26
By observing the relationships between terms.	33

(from page 14)

You think that by representing the main subject by a few terms is characteristic of depth indexing.

Hardly; depth indexing attempts to index a document by *all* the main concepts dealt with. In a research document, many concepts may have to be used as indexing terms; even though they deal in detail with ideas that would not necessarily form part of a brief description of the whole subject of the document.

Now turn to 29, and continue with the course.

(from page 10)

You think that the searcher's needs will be specified precisely if concept terms are translated.

Referring to the authority list may sharpen the searcher's definitions of the concepts in his problem, but this is not the main purpose of this list. He has to know what terms the indexer has used and the authority list offers a standard for them both to refer to.

Now turn back to 10, and read the section on TRANSLATION again. Then select the correct answer.

(from page 11)

You say that pre-coordination of subjects takes place during both indexing and searching.

No. Pre-coordination takes place only during one phase. Conventional indexing assembles terms to make a single description (though perhaps a complex one) of the whole subject of the document. This must obviously be done at the indexing stage; the searcher's job is to recognize the appropriate pre-coordinated description relevant to his inquiry.

Now turn to page 14, and continue with the course.

(from page 19)

It seems to you that by ovserving the rules for combining terms the indexers and searchers can be sure of using the same terms to express indexing concepts.

No. There are no such rules in coordinate indexing. The thesaurus is an authoritative list of terms and the relationships that already exist between these terms. By referring to the list indexers and searchers can be sure that they describe subjects by the same terms.

Now turn back to 19, and read the section on INDEXING LANGUAGE again. Then try the question again.

(from page 29)

You think that combining all concept terms to briefly describe a document characterizes a post-coordinate system.

No. Combination of terms at the indexing stage is characteristic of conventional systems.

Return to 29, read the section on POST-COORDINATION, and consider the question again.

(from page 31)

No, you are wrong. The relationship you have chosen concerns activities and their agents. Teaching is an activity of which teachers are the agents, just as research is an activity of which research workers are the agents. Activities and persons cannot be in genus/species relationship. The genus/species relationship refers only to kinds of things.

Now try again.

Which of the following would be considered a genus/species relationship for the purpose of ERIC indexing?

Students and High School Students	28
Students and Student unions	35

(from page 19)

You think that by referring to the same indexing language, indexers and searchers can be sure of using the same terms to express indexing concepts.

You are right; this is absolutely necessary. The indexing language for ERIC is called a thesaurus.

ERIC THESAURUS

The ERIC Thesaurus is the authority list which serves as a bridge between the language of the searcher and the language of the documents.

The ERIC Thesaurus consists of an alphabetical list of terms with cross references to related terms. It is therefore, in essence, the same as a traditional subject heading list. The Western Reserve Thesaurus is available to give further guidance to ERIC indexers. As well as an alphabetical list of terms it also includes a classified list. This makes it a genuine thesaurus in the original sense of the term.

Which of the following statements best describes the ERIC Thesaurus?

An alphabetical list of indexing terms with cross references to related terms.	30
A thesaurus.	34

(from page 32)

No, you are wrong. The term you have chosen indicates an increase in specificity, since it is one of the members of the group described by the basic term. Retarded reader, for example, is one kind of reader, and is therefore a more specific term. What you are looking for is a term that is broader and indicates a loss of specificity.

Now try again.

Using Information retrieval instead of Indexing.	36
Using Classification instead of Indexing.	39

(from pages 25, 31, or 35)

Right. This is a genus/species relationship, showing that the member of the class is a kind of its class. For example, a professor is a special kind of teacher who teaches in an institute of higher learning.

GENERAL/SPECIAL RELATIONSHIPS

Sometimes one subject is contained by another even though it does not belong to it as a kind. For example, "teaching" belongs to "education" although it is not a kind of "education." Similarly, a more complex subject "the use of television in science teaching in high schools" belongs to a class called "science teaching in high schools" although again it is not a kind of "science teaching in high schools." This sort of relationship is obviously broader than the genus/species relationship we met a moment ago, and to distinguish it we shall call it the GENERAL/SPECIAL RELATIONSHIP.

Which of the following display a general/special relationship?

Chemistry and Chemical apparatus.	32
Chemistry and Physics.	38

(from pages 14, 20)

Yes. If a searcher is to be able to extract very precise information, documents must be indexed in detail. The depth of indexing will be determined by the policy of the individual library.

POST-COORDINATION

Post-coordination is the indexing of documents mainly by isolated terms. Their coordination is left to the searching stage. When the searcher has identified his needs, he can combine terms as he wishes to describe the concepts involved in his study. How the search is made is described later. Post-coordination is the essential difference between conventional and coordinate indexing.

Which statement below best characterizes a post-coordinate indexing system?

Subjects are indexed in depth.	16
Documents are indexed mainly by isolated terms.	19
The indexer combines all concept terms to briefly describe a document completely.	24

(from page 26 or 34)

Good. The ERIC thesaurus consists of an alphabetical list of terms, with related terms. The purpose of the thesaurus is to achieve coincidence of vocabulary, that is, the use of the same terms for given concepts by both the indexer and the searcher.

CONCEPTS AND RELATIONSHIPS

In the classification schedules and in the classified part of a thesaurus, indexing terms are grouped into categories. Relationships between terms are the corollary of these categories. In coordinate indexing the indication of relationships is reduced to a minimum, but for reasons that will become apparent during the course, it is important that they should be clearly and explicitly recognized at the concept indexing stage. There are several (although not many) kinds of relationship and we shall learn about them during the course. First let us look at one of the most important.

GENUS/SPECIES RELATIONSHIP

You may have heard this term used in connection with scientific orders of plants and animals, where it is used strictly to mean a family and its members recognized by a natural fundamental feature. For example, warm blood is a characteristic feature of mammals, no matter where they happen to live, so that a whale is more closely related to other mammals than to fish.

Continue on page 31.

(from page 30)

But this use of the term is too restrictive for library indexing and for our purposes we can include other characteristic features that distinguish kinds of things, like the sex or age of human beings to produce divisions into male and female, or high school and elementary school students.

Remember, however, that the genus/species relationship refers only to kinds of things; do not confuse it with other sorts of relationships or aspects like the properties or processes of things. For example, academic ability is not a kind of student but a property of students; rehabilitation is not a kind of student but an operation on students.

Which of the following would be considered a genus/species relationship for the purpose of ERIC indexing?

Teaching and Teachers.	25
Teachers and Professors.	28
Teachers and one-teacher school.	35

(from page 28)

Good. Chemical apparatus is not a kind of chemistry, but it belongs in the field of Chemistry in what we have called a general/special relationship.

SPECIFICITY

It should now be obvious that the word "specificity" may be used either in the stricter sense in which it implies the species of a genus, or in the looser sense in which it means any class which is contained within another. It is unfortunate that there has been no agreement among writers to use different words for the two ideas, but we hope that when you are aware of the distinction, the context will clarify the meaning intended when the term is used in the text.

Which of the following substitutions of terms indicates a loss of specificity?

Using Pacing in place of Reading.	39
Using Retarded readers in place of Readers	27
Using Language arts in place of Reading.	36

(from page 19)

It seems to you that by observing the relationships between terms, indexers and searchers can be sure of using the same terms to express indexing concepts.

This answer has some relevance, as will be explained later, but it is not the best answer to this question.

The importance of the thesaurus lies in its authority as an agreed list of terms as well as of the relationships between those terms. By referring to this list both indexers and searchers can be sure that they describe subjects by the same terms.

Now turn back to page 19 and read the section on INDEXING LANGUAGE again. Then try the question again.

(from page 26)

You say that the ERIC Thesaurus is best described as a thesaurus.

No. A real thesaurus includes both an alphabetical and a classified listing of terms.

The ERIC Thesaurus, like many coordinate index thesauri, is simply a list of terms in alphabetical order including cross references to related terms.

Now turn to page 30 and continue with the course.

(from page 31)

No, you are wrong. The relationship you have chosen is not a genus/species relationship. A one-teacher school, for example, is not a kind of teacher, but a kind of school; a student union is not a kind of student but a kind of student institution.

Now try again.

Which of the following would be considered a genus/species relationship for the purpose of ERIC indexing?

Research and Research workers.	25
Leadership and Informal leadership.	28

(from pages 27, 32, or 39)

You are correct. For example, Reading is one of several language arts. Other members of the contained classes are Listening, Speaking, and Writing. All are contained within Language arts.

SPECIFICITY IN RELATION TO EXHAUSTIVITY

When we defined DEPTH INDEXING we also called it "exhaustive indexing." Exhaustivity refers to the number of concepts that we recognize for indexing. "Exhaustivity" and "specificity" are the chief characteristics that affect the operating efficiency of an index.

If you leave out a concept (reduce exhaustivity) there is no way of finding the document in response to a search for that concept. If you use a more general term for a concept (reduce specificity) you can find it, though with more difficulty.

Which of the following descriptions reduce specificity without reducing exhaustivity in "Audio-visual aids in language teaching in California"?

Audio-visual aids in language teaching.	42
Audio-visual aids in language teaching in the USA.	46

(from page 46)

You think that concept indexing and translation are parallel activities in indexing and searching.

No. These are consecutive activities in indexing. As we have seen already, concept indexing is the recognition by the indexer of the ideas present in a document; translation is the separate, later activity of expressing those ideas in the agreed language of the thesaurus.

Now return to 46, read the sections on SEARCHING and MATCHING again, and try the question a second time.

(from page 28)

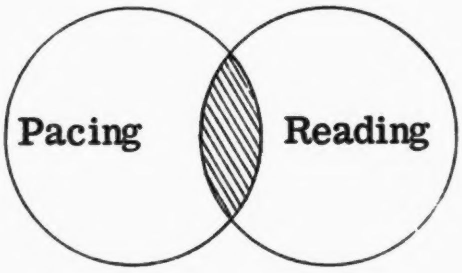
No. Physics is a subject that does not belong in Chemistry at all.

Now go back to page 28, and read the lesson carefully before trying the question again.

(from page 32)

No. Neither of these terms is contained within the other, so the concept of specificity does not apply.

For example, Pacing is not necessarily confined to Reading, nor is Reading necessarily confined to Pacing. A diagram makes this clear.



Here, the shaded intersection shows that a part of Pacing (i.e.) part of the literature dealing with Pacing) refers to Reading and a part of Reading refers to Pacing. But neither Pacing nor Reading is totally contained within the other.

Now try again. Which of the following indicates a loss of specificity?

Using Book indexing instead of Indexing.	27
Using Information retrieval instead of Indexing.	36

(from page 46)

You think that that translation and matching are parallel activites in indexing and searching.

No. Translation takes place both in indexing and searching. Matching takes place only in searching. Translation is the expression in the agreed vocabulary of the thesaurus of ideas present in a document or in an inquiry. Only when translation has been done can inquiries and document descriptions be matched.

Now return to page 46 and read the sections on SEARCHING and MATCHING again. Then try the question a second time.

(from page 44)

No. The words "other than" should tell you that this is a logical negation, that is, a statement that includes *all* of something *but not* a particular part or aspect of it. Logical negation is comparatively rare in indexing and searching; the commonest form is logical product — the simple combination of several terms — and this is what you were looking for here.

Now try the question again.

Which of the following is a logical product?

The Physics of Music.	47
Physics or chemistry.	50

(from page 36)

No. "Audio-visual aids in language teaching" omits the term "California" and thus not only reduces specificity but also exhaustivity because the number of concepts represented has been reduced from three to two.

What you should have chosen was "Audio-visual aids in language teaching in the USA" because although "USA" is not as specific as "California", the inclusion of the concept of place has kept the numbers of concepts at three as in the original statement.

Now turn to page 46 and continue with the course.

You think that Concept indexing and Question analysis are parallel activities.

Yes. Concept indexing is to indexing as question analysis is to searching. The question is made as complete as possible, in the searcher's own words and, as in indexing, the concepts are translated into the controlled language of the system. Matching is the process of finding the terms in a document description that have been stated in the question prescription.

KINDS OF SEARCH

The commonest form of search is one in which a set of terms is asked for in combination. This form is known as the logical product and is symbolized as A x B x C, etc. Teaching science by television is an example of logical product. (Teaching x Science x Television).

There are two less common forms of search which may be used — logical sum and logical negation.

Logical sum refers to a search in which two or more alternatives are acceptable. It is usually referred to as the "and/or" search. An example is a search for documents on either filmstrips and motion pictures or both. It is symbolized as A + B.

Logical negation refers to a search in which some aspect of a subject is unacceptable, for example Audio-visual aids but not Television, it is symbolized as A - B.

Continue on 44.

(from page 43)

Which of the following is a logical product?

Romance languages other than French.	41
Teaching romance languages.	47
French or Italian.	50

(from page 56)

You say that A represents the coordinate index system and B represents the conventional index system.

No. If you are searching for a compound subject in a conventional index system, you will find these subjects ready-made, for example:

A. Conventional Index

1. Dramatic Arts — Creativity — Tests
2. Creativity — Dramatic Arts
   see Dramatic Arts — Creativity
3. Tests — Creativity — Dramatic Arts
   see Dramatic Arts — Creativity — Tests

In the coordinate index system, however, these will be listed as separate terms and you will have to make your own combinations:

B. Coordinate Index

1. Tests
2. Creativity
3. Dramatic Arts

The assumption is that if a searcher wants to know about Tests of Creativy in Dramatic Arts, any documents which have been indexed by the individual terms, Tests, Creativity, and Dramatic Arts, will be relevant to his needs. This is not always true, however, and some of the reasons will be mentioned later in the course. But for simplicity, we will start with the assumption that it is true.

Now you may continue with the course. Please turn to page 59 and go on reading from EXAMPLE OF A COORDINATE INDEX.

(from page 36 or 42)

Good. "Audio-visual aids in language teaching in the USA" keeps the same number of concepts as in the original statement, although it reduces the specificity of one of the concepts — place — by using the more general term "USA" instead of the more specific "California."

SEARCHING

Searching is the operation of matching the inquiry prescription with the document descriptions.

The searching activity is parallel to indexing. The first step is analysis of the question. Just as the main concepts of a document must first be determined and expressed in the indexer's own words, so the searcher's question must be analyzed to determine his real needs. This is parallel to concept indexing, described above, and is similarly followed by translation.

MATCHING

The translated prescription of the question is compared with translated descriptions of the documents. This process is called matching. When the terms used in the question and the document description match, the documents may be expected to be relevant.

Which of the following pairs of terms describe parallel activities in Indexing and Searching?

Concept indexing/Translation	37
Translation/Matching	40
Concept indexing/Question analysis	43

(from pages 41, 44, or 50)

Good. The logical product is the simple combination of the terms in question: Teaching x Romance languages, Physics x Music, or Research x Physics.

EFFICIENCY

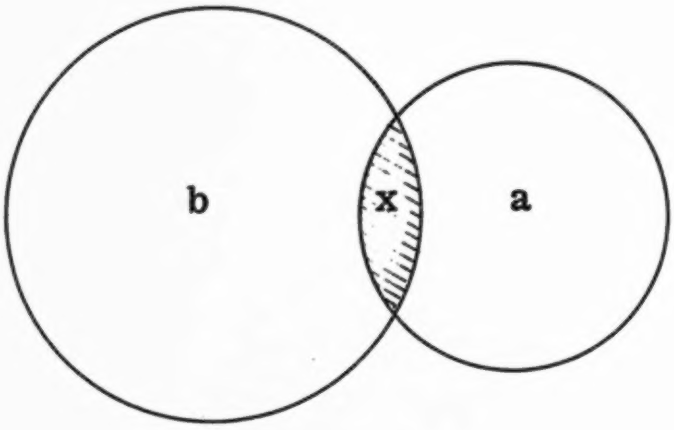
The two criteria of efficiency to be discussed in this course are:

Recall Ratio. Recall Ratio is the number of documents (relevant to an inquiry) retrieved in a search, expressed as a proportion of the total number of relevant documents in the collection.

Precision Ratio. Precision ratio is the number of documents (relevant to an inquiry) retrieved in searching, expressed as a proportion of the total number of documents retrieved.

The following example and diagram may be helpful to illustrate the definitions of Recall Ratio and Precision Ratio.

If there are ten relevant documents (a) in the collection and of these relevant ones seven (x) are retrieved, the Recall Ratio is 7/10 or 70 percent. If 35 documents are retrieved (b) and 7 are relevant (x), the precision ratio is $\frac{7}{35}$ or 20 percent.



circle a = relevant documents

circle b = retrieved documents

x = relevant documents retrieved

$\frac{x}{a}$ = Recall ratio

$\frac{x}{b}$ = Precision ratio

Continue on page 48.

(from page 47)

If x = 5 in the above example (instead of 7) what will the Recall and Precision Ratios be?

Recall, 50%; Precision, 14%	51
Recall, 50%; Precision, 50%	58
Recall, 14%; Precision, 50%	61

(from page 61)

You say that document No. 31 will be relevant to the question.

No. Look again. This number appears on two cards under Teaching and Science but it does not appear under Television. Since this term has not been used in indexing this document, it obviously does not deal with this concept which is important to your search.

Return to 60 and try again.

(from page 44)

No. The word "or" should tell you that this is a logical sum, in which either term or both may be the object of a search. What you were looking for was a logical product (which is the commonest of the three kinds of search discussed in this lesson) and this is simply the combination of several terms.

Now try again.

Which of the following is a logical product?

Physics other than molecular physics	41
Research in physics.	47

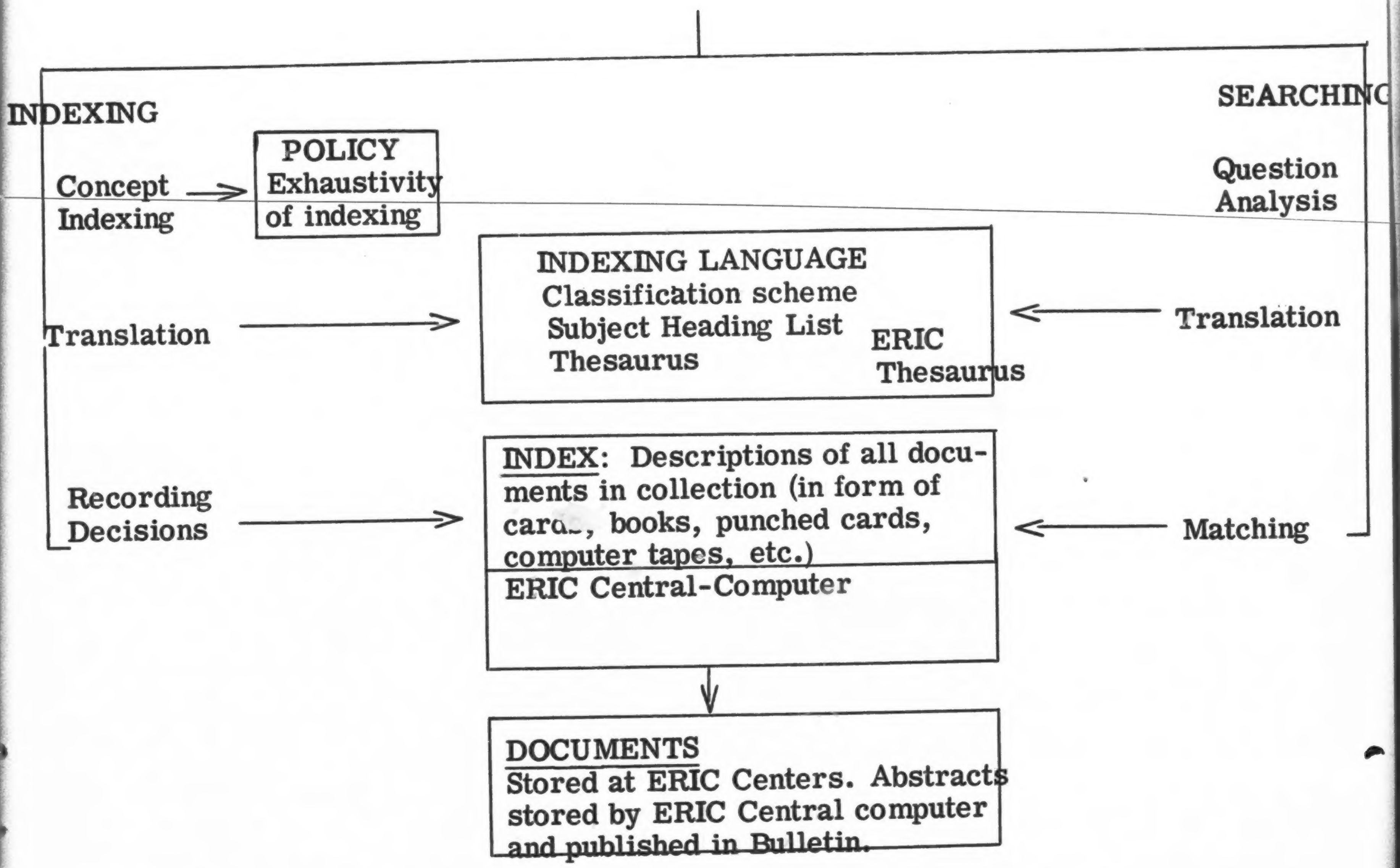
(from page 48)

Your answer: Recall 50%; Precision, 14%

You are correct. Recall Ratio refers to the proportion of relevant documents retrieved from the collection, while Precision Ratio refers to the proportion of relevant documents of those retrieved.

Let us now take a few moments to review a diagram of the major operations of information retrieval.

INFORMATION RETRIEVAL


INDEXING
SEARCHING
Concept Indexing
POLICY Exhaustivity of indexing
Question Analysis
Translation
INDEXING LANGUAGE
Classification scheme
Subject Heading List
Thesaurus
ERIC Thesaurus
Translation
Recording Decisions
INDEX: Descriptions of all documents in collection (in form of cards, books, punched cards, computer tapes, etc.)
ERIC Central-Computer
Matching
DOCUMENTS
Stored at ERIC Centers. Abstracts stored by ERIC Central computer and published in Bulletin.

Continue on page 52.

(from page 51)

The aim of this course is to provide sufficient theoretical explanation and practical demonstration for an indexer to work at an ERIC center. It is necessary, therefore, to understand the main features of its indexing system. These can best be explained by beginning with traditional systems with which you will have had previous experience.

The next section of the course is a brief summary of the conventional index. Read it carefully; you will not be tested on this section but it will give you a useful basis for understanding what comes after it.

THE CONVENTIONAL INDEX

Most libraries use two devices for retrieving relevant documents from the collection.

(1). The systematic order in which documents are arranged by means of a classification scheme.

(2). An index or catalog.

Physical arrangement of documents, however well done, cannot insure that all documents relevant to an inquiry are collected at one point. The reasons may be classified as physical and intelluctual.

(1). Physical reasons:

Information may be in different forms (books, periodicals, phonograph records), in different sizes, in different departments for administrative convenience, and many items may be out on loan.

(2). Intellectual Reasons:

Inquiries for documents by a given author and for documents on a given subject cannot be met by the same arrangement.

Continue on the next page.

(from page 52)

It is not practicable to multiply copies of documents for this purpose, but it *is* practicable to multiply index entries which *represent* the documents. Thus, the minimum provision in an index is two entries for each document — one for author and one for subject.

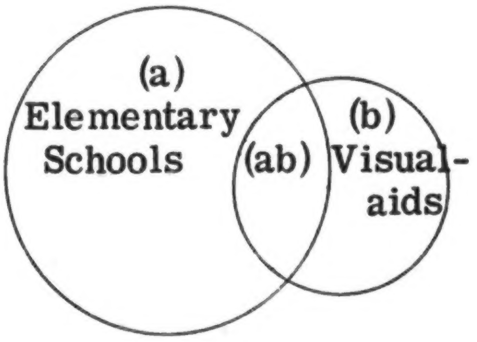
From now on we concentrate on the problem of subject entries. It is important to bear in mind that in most libraries the indexing consists of summarization. Thus it is common for a large document to be represented by a brief subject description, such as Education or Universities. Documents on either of these subjects may deal with hundreds of concepts, but by and large they are all *implied* by the summary. The only way to be certain whether a specific part of either subject is actually dealt with in the work concerned, is to consult its own index, which should be an example of an exhaustive (or depth) index.

Both Education and Universities are single concepts represented by single words. Elementary schools is also a single concept, but represented by two words. The indexer has the choice of using either *Elementary schools* or *Schools, elementary* as the heading for this concept. There are several factors to be considered in making the choice, but the most important thing is that a cross reference should be provided from the rejected form.

In all of the above cases it is possible to keep all entries relating to a subject in one place. But with a subject like Visual-aids in elementary schools, the problem is somewhat different. Here we no longer have a single concept. Consider the subjects as shown in the diagram on the next page.

(from page 53)

If the entry is made under Visual-aids (or Aids, visual) the set of entries relating to Elementary schools is deprived of some of its members; if we make the entry under Elementary schools (or Schools, elementary), then the Visual-aids class is similarly incomplete.



ab = Visual-aids in elementary schools and clearly belongs to both class (a) and class (b).

There are three ways in which this problem may be treated.

(1). Single entry with cross reference. If the library chooses the heading Elementary schools — Visual-aids, then a reference must be made:

Visual-aids - Elementary schools see Elementary schools — Visual-aids. This insures that the person interested in visual-aids will know that there are entries relating to his subject in other parts of the catalogue.

The advantage of the method is in economy. The disadvantage is that the user has to look in more than one place for his entries.

(2). Multiple entry. The library makes entries under both Elementary schools — Visual education and Visual education — elementary schools. The advantage of this method is that there will be a complete set of entries filed under every term. For example, the person interested in Visual-aids will find at one point entries relating to every aspect of the subject.

The disadvantage is that it causes an increase in the size of the index.

(3). Non-specific multiple entry. In (1) and (2) the entries and references consist of complete statements of the subject. In most library catalogues in the United States multiple entries are provided, but each entry is only a partial statement of the subject. A similar method is adopted by the Education Index. For example, a document on Counseling the adolescent girl on personal problems has the entries.

Continue on the following page.

(from page 54)

Girls.
Parent-child relationship.
Personal problems—High school students.
Adolescence, see also Girls.

The advantage of this method is that it avoids the difficult problems of term order in complicated headings. The disadvantage is that the precision ratio suffers seriously. A question on "Counseling adolescent girls" would be retrieved only by examining all the items on Girls (alone) or on Adolescents (alone).

This brief summary of practice makes clear that it is difficult to provide at one and the same time for the four major requirements of a specialized research collection:

(1). Detailed indexing, which is necessary for high recall ratio.
(2). Specific descriptions, which are necessary for high precision ratio.
(3). Equally easy access to all aspects of a subject.
(4). Reasonable size of index.

Coordinate indexing arose as an answer to the problem. We need not concern ourselves here with its limitations or with alternative answers. The remainder of the course will concentrate on the technique of coordinate indexing in general and its application to ERIC in particular. A bibliography is given at the end of the course for those who wish to study the subject further.

Please continue on 56.

(from page 55)

COORDINATE INDEXING THEORY

Compound subjects create one of the main difficulties in indexing. In the conventional index, compound subjects are represented by phrases; the coordinate index represents them as separate terms.

The corollary of indexing is searching. In a conventional index, the searcher finds his compound subjects ready-made. In a coordinate index, the searcher has to make his own combination of terms.

Here is a document that presents this problem: "Tests of Creativity in the Dramatic Arts."

A and B below represent the two different methods of indexing this document.

A	B
1. Dramatic Arts—Creativity—Tests	1. Tests
2. Creativity—Dramatic Arts *see* Dramatic Arts—Creativity	2. Creativity
3. Tests—Creativity—Dramatic Arts *see* Dramatic Arts—Creativity—Tests	3. Dramatic Arts

Which of the following is the correct statement?

A represents the coordinate index system;
B represents the conventional index system. 45

A represents the conventional index system;
B represents the coordinate index system. 59

(from page 60)

You say that document No. 20 will be relevant to the question.

No. Look again. This number appears on two cards, under the terms Teaching and Television. But it does not appear under the third term, Science. Since this term has not been used in indexing this document, the document obviously does not deal with this concept which is important to your search.

Return to 60 and try again.

(from page 48)

Your answer is Recall, 50%; Precision 50%

Hardly; if we assume that there are 10 relevant documents in the collection and that five of these are retrieved, then it is true that the Recall Ratio will be fifty percent. But we also assumed that altogether 35 documents were retrieved, not ten. Remember, Precision Ratio refers to the number of documents actually retrieved, not to the number in the collection.

Go back to 48 please, recalculate, and select the right answer.

(from page 45 or 56)

Yes. If you are searching for a compound subject in a conventional index system, you will find these subjects ready-made (pre-coordinated) as in Column A. In the coordinate index system, however, you will have to make your own combination of the single terms represented in Column B (post-coordination). The assumption is that if a searcher wants to know about Tests of Creativity in Dramatic Arts, any document which has been indexed by the individual terms Tests, Creativity, and Dramatic Arts will be relevant to his needs. This is not always true, however, and some of the reasons will be mentioned later in the course. But for simplicity, we will start with the assumption that it is true.

Example Of A Coordinate Index

The coordinate index may be handled by a variety of devices, ranging from hand-written cards to computers. The principle of operation is always the same. One of the simplest forms in use is the Uniterm Index. Each subject in the system is represented by a card on which the numbers of the documents indexed are recorded. Numbers are arranged in columns by their last digit to facilitate matching.

If the document on Tests for Creativity in Dramatic Arts were number 31, the entries might appear as follows:

Tests			
20	31	32	
	51		

Creativity			
20	31		
	41		

Dramatic Arts			
40	31		23

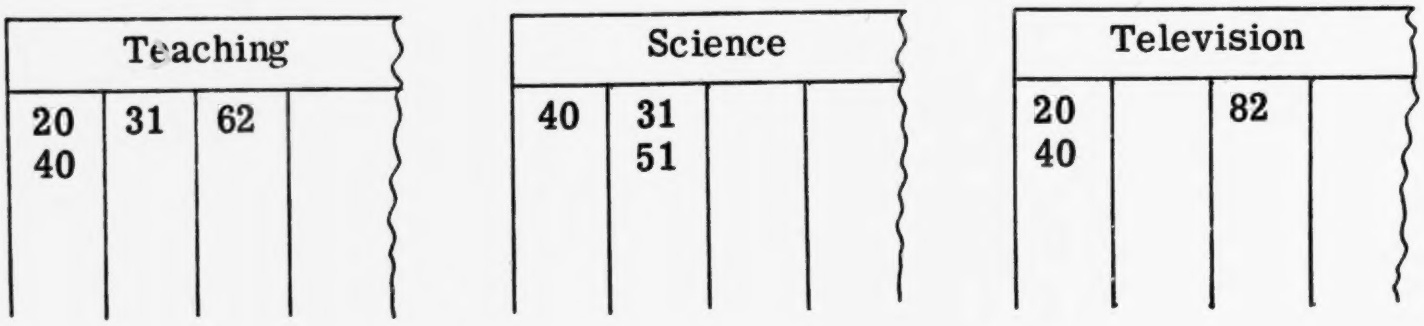
The searcher looking for documents on Tests of Creativity in Dramatic Arts would take these three cards from the alphabetically-arranged file and compare them. Any number appearing on all three cards would indicate documents to be examined. In this case, document No. 31 appears under all three terms.

Continue on 60.

(from page 59)

A perfect match is achieved when the indexer and searcher use precisely the same words to describe a subject. We have seen how a thesaurus can facilitate this matching by controlling the choice of individual terms.

Suppose the search is for a document on Teaching Science by television and you examine the following cards:


Teaching
20
40
31
62
Science
40
31
51
Television
20
40
82

What is the number of the document that will be relevant to this search?

Document No. 31	49
Document No. 20	57
Document No. 40	63

(from page 48)

Your answer: Recall 14%; Precision 50%

Definitely wrong. We are still assuming that there are ten relevant documents in the collection. For purposes of the problem, we said that five of these were retrieved. If the Recall Ratio is related to the number of relevant documents in the collection, the ratio could not be $\frac{5}{35}$. The Precision Ratio refers to the number of documents retrieved, not to the number in the collection.

Go back to 48 please, recalculate, and select the right answer.

(from page 68)

No, the document you have selected is certainly the one that appears only under the terms in question, but you should also have more documents — those that appear not only under those terms, but under any others as well.

Now try again.

If the searcher looked in the index for only Physics and Teaching, which group of documents would he retrieve?

Physics		
10	21	22
	31	32
		42

Teaching		
30	11	22
40	31	32
50		42

Television		
10	31	22

High School		
10	31	

22, 32, 42	71
32, 42	74
22, 31, 32, 42	82

(from page 60)

You say that document No. 40 will be relevant to the question.

Right, since document No. 40 is the only one that appears on each of the cards.

In all indexing, the objective is to find a meeting point between authors with something to teach and readers with something to learn. The problem is variety of language. Very similar documents may use very different terminology; identical questions may be phrased in a variety of ways. The indexer must be consistent in always using the same word (or words) to represent an idea. The searcher must use the same words as the indexer.

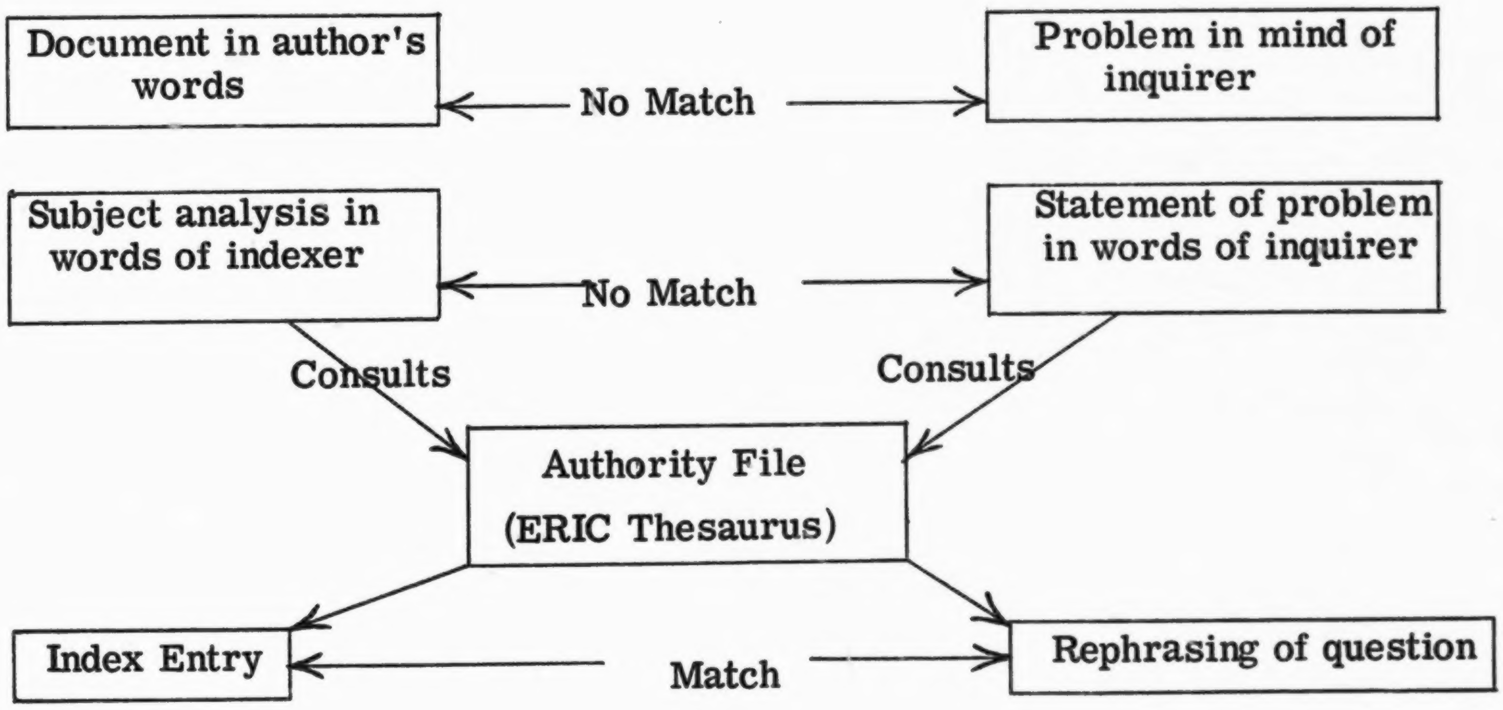
How can this be accomplished?

Indexers must agree on the terms used to represent ideas.	66
An authority list must be set up for indexers and searchers.	69
Authors must agree on the same terms to represent concepts.	72

(from page 67 or 70)

Most people would regard TEACHING and INSTRUCTION as synonyms, but in the specialized context of the ERIC system they are used as separate terms representing different aspects of the topic. You must be very careful to decide when and why you distinguish between terms. If you do not do this, you may either use the same term for different aspects of a subject (when you should have used the other term) or you may index related or identical subjects under different terms.

The first job of ERICT then, is to show which words may be used for indexing and searching. To index a particular concept, the indexer consults ERICT to find which word or term he may use to index the document. To search for material on a particular concept, the inquirer also consults ERICT to see which word or term he should use in his question. As a result the indexer and the inquirer are both using the same language — the language of ERICT. This provides an efficient route to the documents relevant to the inquirer's needs. The following diagram illustrates the process.


Document in author's words
No Match
Problem in mind of inquirer
Subject analysis in words of indexer
No Match
Statement of problem in words of inquirer
Consults
Consults
Authority File (ERIC Thesaurus)
Index Entry
Match
Rephrasing of question

Please continue on 65.

(from page 64)

Other Variations Between Index Description and Search Prescription

A perfect match is achieved when indexer and searcher use precisely the same terms to describe a subject. We have seen how ERICT can facilitate this matching by controlling the choice of individual terms.

There are several other ways in which the indexing description and search prescription may vary. One is the variation in number of terms used and another is in the variation in generic level of terms (whether we specify the genus or any of its species).

(1). Variation in number of terms

An inquirer asks for Creativity in Dramatic Arts. There are two possibilities:

(a). He really wants to see everything which discusses this topic.

(b). He really wants a narrower subject, but has failed to state it precisely.

Let us suppose that the searcher wants to see everything which discusses creativity in dramatic arts. If he uses only the terms "Creativity" and "Dramatic Arts" in his search question, will he be able to retrieve all the documents that include "Creativity" and "Dramatic Arts" in their indexing?

Yes	68
No	76

(from page 63)

You say this can be accomplished by having the indexers agree on the terms used to represent ideas.

This is only part of the answer. Even if the indexers use the same terms, we still have the problem of the searchers. Unless they know what terms have been used by the indexers, there will still be problems in retrieval of relevant documents.

Please return to 63 and find another solution.

(from page 70)

You would treat TEACHING and INSTRUCTION as separate terms – so does the ERICT. Most people would regard TEACHING and INSTRUCTION as synonyms, but in the specialized context of the ERIC system they are used as separate terms representing different aspects of the topic. You must be very careful to decide when and why you distinguish between terms. If you do not do this, you may either use the same term for different aspects of a subject (when you should have used the other term) or you may index related or identical subjects under different terms.

Now continue with the lesson on 64.

(from page 65 or 76)

You think that if the search includes only the terms "Creativity" and "Dramatic Arts", the inquirer *will* be able to retrieve all the documents that include these two terms in their indexing.

Yes. If an inquirer really wants to see *everything* which discusses the subject of creativity in dramatic arts, he is well served by coordinate indexing. A search by these two terms (Creativity and Dramatic Arts) will produce all the documents indexed by them, *irrespective of any additional terms by which they are indexed*. If we use *A* to stand for "Creativity", and *B* for "Dramatic Arts", the search for A x B will not only produce documents indexed solely by these terms. It will also produce documents indexed by A x B x C x D, etc. A document indexed by A (Creativity), B (Dramatic Arts), and C (Testing), *would be retrieved* in this search and *could be relevant* to the inquirer's needs.

If the searcher looked in the index for only Creativity and Dramatic Arts, which group of documents would he retrieve?

Creativity			
20	31	32	
60	41	52	
	71		

Dramatic Arts			
40	31	32	
	71	52	

Testing			
20	31	52	
	51		

Girls			
60	71	52	

32	62
31, 32, 52	71
31, 32, 52, 71	82

(from page 63)

You say this can be accomplished by having an authority list for indexers and searchers.

Right. Not only must the indexers for ERIC be consistent in always using the same word or term to represent a concept, but the searcher must have some way of knowing what word or term the indexer has used to represent that concept. The only way this can be accomplished is by the use of an authority list which will suggest one term to represent a particular concept. In our case, this authority list is the ERIC Thesaurus (ERICT).

Please continue on 70.

(from page 69)

SYNONYMS AND VARIANT WORD FORMS

The simplest variations in language are the result of synonyms and varying grammatical forms of the same root. In his inquiry for documents on the subject of Tests of Creativity in the Dramatic Arts, the searcher may use Theater as a synonym for Dramatic Arts, or he may use Testing instead of Tests and thus fail to find useful information.

The authority list can facilitate the matching process by choosing one form in each case and referring from the other. For example:

Theatre
Use DRAMATIC ARTS

Testing
Use TESTS

DRAMATIC ARTS
UF (i.e. Use for) Theatre

TESTS
UF (i.e. Use for) Testing

In this way ERICT can direct all documents and all searches for different versions or expressions of the same idea to a single place. "Synonym" is a relative term — very few pairs of words have precisely the same meaning. Very often it is the job of ERICT to tell the indexer and the searcher that for practical purposes they need not distinguish between words with closely related meanings.

On the other hand, ERICT sometimes includes apparent synonyms and variant word forms. This may be because each of a pair of words that are loosely synonymous in everyday speech can be used to represent two aspects or uses of a term that must be distinguished in the context of the special field for accurate subject description and economical searching.

If you were constructing a thesaurus for education would you treat Teaching and Instruction as synonyms?

Yes 64

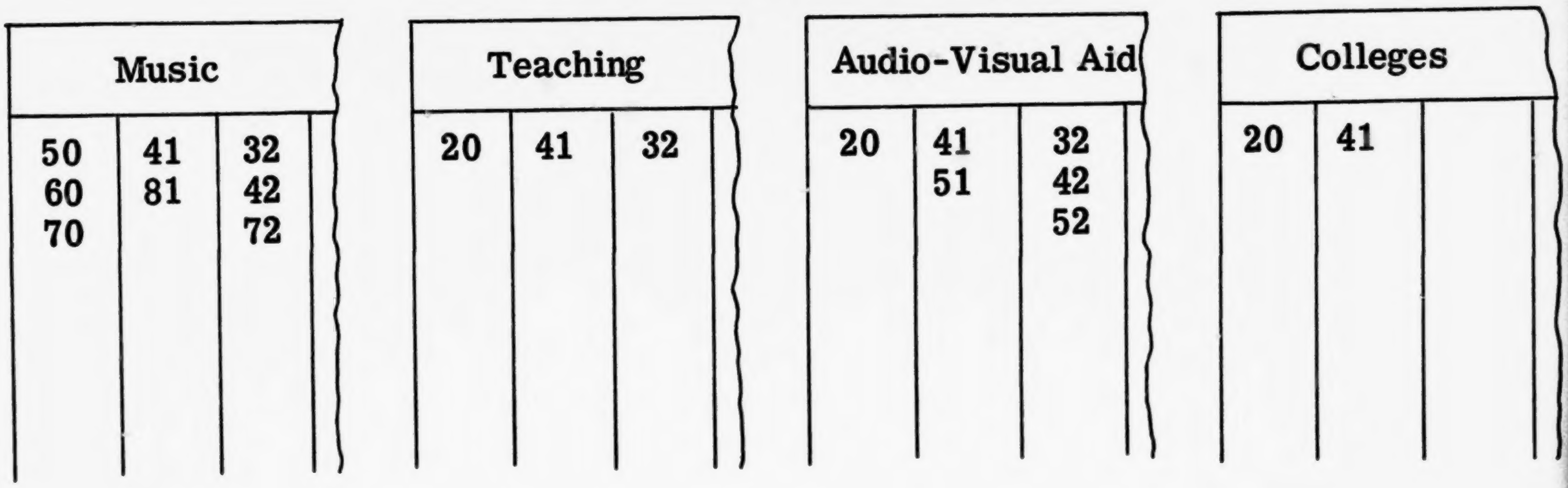
No 67

(from page 68)

No, although you are almost right. You have listed not only the documents that appear only under the terms in question, but also documents that appear there and under other terms as well. But you have not listed all of them.

Now try again.

If the searcher looked in the index for only Music and Audio-visual aids, which group of documents would he retrieve?



Music			
50 60 70	41 81	32 42 72	

Teaching			
20	41	32	

Audio-Visual Aid			
20	41 51	32 42 52	

Colleges			
20	41		

42	62
32, 42	74
32, 41, 42	82

(from page 63)

You say this can be accomplished by having authors agree on the same terms to represent concepts.

This is impossible. And this is exactly why we need to find some system to make the retrieval of documents possible. Similar documents may use very different terminology and searchers may ask identical questions phrased in a variety of ways.

Return to 63, and find another solution to this problem.

(from page 82 or 85)

Good. Both these documents concern the three terms in question, even though one of the documents actually involves a fourth term as well. If you wished, you could narrow the search even further by adding a fourth term, and thus retrieve the single document — but the point has now been made sufficiently.

As you have seen, a search for Creativity x Dramatic Arts x Testing will not retrieve a document indexed by only Creativity x Dramatic Arts, or by Creativity x Dramatic Arts x Girls. All terms of the inquiry must be included in the index description to insure retrieval of a document.

How can we retrieve a document indexed by Creativity and Dramatic Arts? By searching for

Creativity alone, or Dramatic Arts alone or Creativity x Dramatic Arts. 77

Creativity x Dramatic Arts x Testing, or Creativity x Testing, or Dramatic Arts x Testing. 83

(from page 62 or 71)

No, you have just made the same mistake again.

Coordinate indexing uses terms in isolation, and a search for two terms should reveal all the documents under those terms, irrespective of any other terms used to index them.

Now go back to 68, read the lesson carefully, and try the question again.

(from page 78)

No. To say that School library has a generic relationship to School is to say that it is a kind of School.

Now go back to 78, and read the lesson again, paying particular attention to the list of kinds of non-generic relationships. Then try the question again.

(from page 65)

You think that if the search includes only the terms "Creativity" and "Dramatic Arts," the inquirer will *not* be able to retrieve all the documents that include these two terms in their indexing.

This is not true. If we use *A* to stand for "Creativity" and *B* for "Dramatic Arts," the search for A x B will not only produce documents indexed solely by these terms. It will also produce documents indexed by A (Creativity), B (Dramatic Arts), and C (Testing) *would* be retrieved in this search and *could* be relevant to the inquirer's needs.

Now you may continue on 68.

(from page 73 or 83)

You say a document indexed by two terms, A & B, will be retrieved by searching for A alone, B alone, or A x B.

Yes. A document indexed by A (Creativity) and B (Dramatic Arts) will be retrieved by searching for

A (Creativity) or B (Dramatic Arts) or A x B (Creativity x Dramatic Arts)

But it will not be retrieved by searching for

A x B x C (Creativity x Dramatic Arts x Testing) or A x C (Creativity x Testing) or B x C (Dramatic Arts x Testing)

VARIATION IN GENERIC LEVEL OF TERMS

Subjects may be varied by the generic level of individual terms as well as by variations in the number of terms. For example, instead of Creativity in Dramatic Arts, we may think of Ability in Dramatic Arts. In this case, the generic level of Creativity has been raised, to that of the broader term Ability. Or we may think of Creativity in Acting. In this case the generic level of Dramatic Arts has been lowered to that of the narrower term Acting.

Neither a search for Ability x Dramatic Arts, nor a search for Creativity x Acting would produce a document indexed by Creativity and Dramatic Arts, though it may be relevant in either case.

The searcher must be aware of generic relations, so that he may either look for more material related to his inquiry, or make his prescription more precise. These procedures are discussed in Lesson 2.

Continue on 78.

(from page 77)

Other Variations

Apart from generic relationships there are others, such as part, property, process, operation. For example:

Film is a kind of audio-visual aid(= Generic Relationship).
Frame is a part of a film.
Color is a property of a film
Projection is an operation in showing a film
Deterioration is a process in a film.

What is the relationship of "School library" to "School"?

Generic	75
Part	81
Property	86

(from page 81)

No. "The Recognition of Synonyms" is a Recall Device that broadens the subject description.

Go back to 81, and read the lesson carefully, then try the question again.

80
(from page 85)

No, you are still making the same mistake. You must always list all the documents indexed under the terms in question, whether they are indexed elsewhere or not. If you wished to narrow the search further you could, of course, search for Audio-visual aids for Teaching Music in Colleges and retrieve a single document (41).

Now go back to 82 and read the lesson carefully. Then try the question again.

(from pages 78, 87)

Good. The School Library is neither a *process* in the School, nor a *kind* of School — it is, as you have said, a *part* of the School.

INDEXING DEVICES

We have learned already that the RECALL RATIO is the number of documents retrieved relevant to an inquiry expressed as a proportion of all the relevant documents in the collection.

Indexing and searching procedures designed to allow us to *broaden* the description of a subject and increase the number of documents retrieved are known as "*Recall Devices*."

Recall Devices

1. *Recognition of synonyms*. (Usually done by the indexer).
2. *Recognition of variant forms of the same root*. (Sometimes done in indexing, sometimes in searching).
3. *Recognition of generically related terms*. (Done in the search procedure).
4. *Recognition of other relationships*. (Done in the search procedure).

All of these recall devices are indexing and searching procedures that broaden the description of the subject and increase the number of documents retrieved.

Which of the following devices would *not* broaden the subject description?

Recognition of synonyms.	79
Recognition of coordination of terms.	84
Recognition of generically related terms.	87

(from page 68)

Good. You have listed all the documents entered under the two terms in question, even though some of them appear under other terms as well.

What if the inquirer really wants a narrower subject, but has failed to state it precisely? In this case, he will see that many of the documents produced are not relevant to his particular need. He may now realize that it is only Testing for Creativity in Dramatic Arts that he wants. The search may now be adjusted to produce only those documents relevant to his needs.

Here are the cards. What documents will fill this last requirement?

Creativity			
20	31	32	
60	41	52	
	71		

Dramatic Arts			
40	31	32	
	71	52	

Testing			
20	31	52	
	51		

Girls			
60	71	52	

31, 52	73
31	85

(from page 73)

No. The document must be indexed by all of the terms in the inquiry before it can be retrieved. In this case, the first version of the inquiry asks for Testing, as well as Creativity and the Dramatic Arts, and it prevents the retrieval of a document indexed simply by Creativity and Dramatic Arts. Similarly the other two versions of the search ask for terms not indexed in the document. The search must have the same, or a less number of terms than the index description before it can be broad enough to retrieve a document with that description.

Now turn to 77, and continue with the course.

(from page 81)

Good. "Coordination of terms" is not a Recall Device that broadens the description of a subject — in fact it narrows it, and that is what we shall look at next.

We have learned that the PRECISION RATIO is the proportion of relevant documents in the number of documents retrieved.

Indexing and searching procedures that are designed to allow us to narrow the description of a subject and reduce the number of documents retrieved (thus increasing their chance of being relevant) are known as precision devices.

Precision Devices

By far the most important precision devices are:

1. Coordination of terms. This is done mainly by the searcher, though there are some ready-made coordinations in ERICT.

Of less importance is:

2. Weighting. The ERIC indexer selects a maximum of five terms to show the most important theme(s) of the document. If an inquiry specifies "only those documents in which x is a major subject," the weighting device allows the search to be selective in this way.

The use of the theory explained in this lesson is demonstrated in Lessons II and III. More detailed accounts are to be found in the books suggested for further reading.

Turn to 89.

(from page 82)

No. The single document you have chosen is certainly the only one to appear only under the three terms in question, but what about the document that appears under all four — it still covers the original three.

Now try again.

What documents cover Teaching of Music in Colleges?

Audio-Visual Aids			
50	41	32	
60	81	42	
70		52	

Music			
20	41	32	

Colleges			
20	41	32	
	51	42	
		52	

Teaching			
20	41		

20, 41	73
20	80

(from page 78)

No. To say that the School Library is a property of the School is to misinterpret the word "property". If you read the lesson on page 78 carefully, you will realize that the word "property" really means "quality".

Now try the question on 78 again.

(from page 81)

No. "The recognition of variant forms" is a Recall Device that broadens the subject description.

Go back to 81, and read the lesson carefully. Then try the question again.

(from page 92)

YOUR ANSWER: The two stages of ERIC indexing are the pre-coordination and the translation stages.

No. You should review this section. Please return to 6 and follow the course until you get to 15. When you reach 15, do not continue with the course, but return to 92, and try this question again.

(from page 84)

QUIZ

Lesson 1

Be sure you consider the answer choices carefully. If you select the wrong answer, you may be asked to return to that part of the program and review the information.

Question 1.

Which of the following is a correct statement?

In conventional indexing the terms are pre-coordinated; In coordinate indexing, the searcher must coordinate the terms.	92
In coordinate indexing, the terms are pre-coordinated; In conventional indexing, the searcher must coordinate the terms.	96

(from page 91)

YOUR ANSWER: The ERIC Thesaurus is best described as a classified list of indexing terms and their synonyms.

No. This may be used to describe Roget's Thesaurus, but not the ERIC Thesaurus. Please return to 26 and review the material on the ERIC Thesaurus. Then return to 93 and continue with this Quiz.

(from page 101)

YOUR ANSWER: Using baroque music instead of music is an example of specificity.

Yes. Since baroque music is a species of the genus music, we would use the more precise term to increase the specificity of indexing.

Question 7.

The term that refers to the number of concepts recognized for indexing is called

Specificity	100
Exhaustivity	105
Search prescription	108

(from page 89)

YOUR ANSWER: In conventional indexing, the terms are precoordinated.

In coordinate indexing, the searcher must coordinate the terms.

Right. Now you may go on to the next question.

Question 2.

What are the two stages of ERIC indexing?

The pre-coordination and translation stages.	88
The concept indexing and translation stages.	95
The indexing and the searching stages.	98

(from page 99)

YOUR ANSWER: The ERIC Thesaurus is best described as a list of indexing terms with cross references to related terms.

Yes. You may continue with your next question.

Question 5.

Which of the following could not be considered a genus/species relationship?

School; Junior High School	97
Teacher; Team teaching	101
Reading; Remedial Reading	104

(from page 105)

YOUR ANSWER: When the translated prescription of the question is compared with the translated description of the document, this part of the searching process is called matching.

Yes. Please go on with the next question.

Question 9.

Logical product, logical sum, and logical negation are three forms of

Indexing	103
Searching	106

(from page 92)

YOUR ANSWER: The two stages of ERIC indexing are the concept indexing and the translation stages.

Right. Continue with your next question.

Question 3.

When the summarized subject involves more than one concept, the combination of terms made when the document is indexed is called

Pre-Coordination	99
Coordination	107
Post-coordination	110

(from page 89)

YOUR ANSWER: In coordinate indexing, the terms are precoordinated. In conventional indexing, the searcher must coordinate the terms.

No. You need a little review. Please return to 52, and follow the course until you get to 59. When you reach 59, do not continue reading that part of the course, but turn back to 89, and try this question again.

(from page 93)

YOUR ANSWER: School and Junior High School could not be considered a genus/species relationship.

Yes, it could. School is the genus of which a particular kind of school (Junior High School) is a species.

Please return to 93 and try again.

(from page 92)

YOUR ANSWER: The two stages of ERIC indexing are the indexing and searching stages.

No. You need some review. Please return to 6 and follow the course until you get to 15. When you reach 15, do not continue with the course, but return to 92, and try the question again.

(from page 95)

YOUR ANSWER: The combination of terms made when the document is indexed is called pre-coordination.

Right. And when the combination is made during the searching stage, it is called post-coordination. Now you may continue with your quiz.

Question 4.

Which of the following best describes the ERIC Thesaurus?

A classified list of indexing terms and their synonyms.	90
A list of indexing terms with cross references to related terms.	93

(from page 91)

YOUR ANSWER: Specificity refers to the number of concepts recognized for indexing.

No. Please return to 36 and review the material on that page. Then return to 91 and try the question again.

(from page 93)

YOUR Answer: Teacher and Team Teaching could not be a genus/species relationship.

Right. Team teaching is a kind of action performed by the agent (teacher). If the words had been Teaching/Team Teaching, there would be a genus/species relationship because team teaching would be a species of the genus, teaching.

Please continue with the next question.

Question 6.

Which of the following illustrates specificity?

Using baroque music instead of music.	91
Using music instead of baroque music.	111

(from page 105)

YOUR ANSWER: When the translated prescription of the question is compared with the translated description of the document, this part of the searching process is called translation.

No. The translation has already taken place. Please return to 46 and review the material on that page. Then return to 105 and try the question again.

YOUR ANSWER: Logical product, logical sum, and logical negation are three forms of indexing.

No. Please return to 43 and review the material on that page. Then come back to 94 and try the question again.

(from page 93)

YOUR ANSWER: Reading and remedial reading could not be considered a genus/species relationship.

Yes it could. Reading is the genus of which remedial reading is a kind of species.

Please return to 93 and try again.

(from page 91)

YOUR ANSWER: Exhaustivity refers to the number of concepts recognized for indexing.

Right.

Question 8.

When the translated prescription of the question is compared with the translated description of the document this part of the searching process is called

Matching	94
Translation	102
Retrieval	112

(from page 94)

YOUR ANSWER: Logical product, logical sum, and logical negation are three kinds of search.

Right. Now for your last question.

Question 10.

Recognition of synonyms, variant word forms, generically-related terms, and other relationships are devices used to increase

Recall	113
Precision	109

YOUR ANSWER: The combination of terms made when the document is indexed is called coordination.

Not quite. We are looking for the term that means the coordination takes place in the indexing stage. Please return to 95 and try again.

(from page 91)

YOUR ANSWER: Search prescription refers to the number of concepts recognized for indexing.

No. The prescription has nothing to do with the number of concepts — only with the terms used to represent the concepts.

Please return to 36 and review the material on that page. Then return to 91 and try the question again.

(from page 106)

YOUR ANSWER: You say that these are devices used to increase precision.

No. The devices used for precision are coordination of terms and weighting. You should go back to 81 and review the material. Then return to 106 and try the question again.

(from page 95)

YOUR ANSWER: The combination of terms made when the document is indexed is called post-coordination.

No. You need to review this section. Please return to 11 and follow the course until you get to 19. When you reach 19, do not go on with the course, but return to 95 and try the question again.

(from page 101)

YOUR ANSWER: Using music instead of baroque music is an example of specificity.

No. You should go back to 32 and review the section on Specificity. Continue with the program until you get to 46, then return to 101 and try the question again.

(from page 105)

YOUR ANSWER: When the translated prescription of the question is compared with the translated description of the document, this part of the searching process is called retrieval.

Not quite. This process leads to retrieval, however. Please return to 105 and try again.

(from page 106)

YOUR ANSWER: You say that these are devices used to increase recall.

Right. For higher precision, we use coordination of terms and weighting.

We hope you have gone through this quiz without having been referred back to the program. But if you were asked to review, we hope that you did so and that the questions you missed are now clear in your mind.

You are now ready to go on to your next lesson.

Lesson 2

HOW TO INDEX A DOCUMENT

INTRODUCTION

The purpose of indexing a research document is to make sure that it will be found if it is relevant to a question put to an ERIC center. Indexing is designed to assist retrieval.

In essence, indexing consists of assigning to the document a number of labels in the form of terms (single words or brief phrases). Together, these terms add up to a comprehensive statement of what the document is about.

The ultimate subject of most documents we deal with in ERIC is education. Even though a document is largely about some psychological problem or some statistical technique, the reason it is handled by an ERIC center is that it contributes to the subject of education, directly or indirectly.

The ERIC indexer must also bear in mind that he is indexing for the whole ERIC system and not just for his center. He should avoid giving preference to his own specialty by choosing only those aspects of the document, or index terms, in which he is interested or which have particular meaning for him.

Now we can begin to learn how to index a document. Turn to 2.

(from page 1)

In indexing for ERIC there are two quite distinct operations. One is concept indexing — determining the significant concepts in a document, and representing them by any suitable terms. The easiest method is to use the terminology of the document.

The second step is the translation — stating these concepts in the language of the ERIC Thesaurus (which from now we shall call ERICT). For example, you may decide that the notion of "Home-school relationship" is a concept to be indexed. During the translation stage you consult ERICT and find that "Home-school relationship" is expressed as FAMILY-SCHOOL RELATIONSHIP. This then, is the translated term with which you will index the document.

How would you define translation?

The use of ERICT to identify significant concepts in a document.	5
The conversion of the significant concepts into the language of ERIC.	7

(from pages 11, 7)

You feel that the selection of candidate terms is best done by using ERICT as you go along.

No. This would certainly interrupt your thoughts and attention. It is important that indexing be done with the least amount of time consistent with accuracy. In any case you cannot be certain of finding all the relevant and correct terms in this way. It is far better first to list all the significant concepts in the document's own language and only then to translate them into the language of ERIC.

When you are selecting candidate terms, is it a good idea to use terms most likely to be found in ERICT?

Yes	11
No	9

4
(from page 9)

You feel that a document should first be scanned, then read thoroughly.

No. This is a good method for the purpose of studying, but we are discussing reading for indexing purposes. We are concerned with the amount of time spent on each document, as well as the quality of indexing. The balance of reading and rapid scanning will vary according to the content of the document and your knowledge and experience. Some documents may be indexed after rapid scanning; others may require careful reading, at least of some sections.

Now turn to 6 and continue with the lesson.

(from page 2)

No. The use of ERICT is not to help you recognize significant concepts but to translate terms you have already chosen at the concept indexing stage. This distinction is important since there can be considerable difference between the terms in a document and the way in which the concepts they represent are ultimately expressed for the ERIC indexes. Only by working independently of ERICT at first can you be certain of having an objective view of the concepts contained in a document.

We shall now study the concept indexing stage in more detail; later we shall deal with the translation process.

Now turn to 7 and continue with the lesson.

(from page 9)

You feel that the balance of reading and rapid scanning will vary according to the content of the document.

Yes; rapid scanning of some documents will provide all the cues that you need for indexing them, while intensive reading will be required to answer your questions about other documents.

Distinguishing Between Form Words and Subject Words

During the first examination, the indexer should decide what sort of document he is dealing with. It is essential to distinguish between two kinds of words that may be used to describe a document — form words or subject words. The majority of terms are subject words — words that tell us what the document is about. There are, however, a few very important words which, though they look like subject words, tell us instead what sort of document we are dealing with.

The simplest example of one word used in two ways is Bibliography. This may indicate that a document is in the form of a bibliography (in which case bibliography would be a form word), or it may indicate that the document is about bibliography, for example, how to make a bibliography or how to teach the subject (in which case bibliography would be a subject word).

Why is it necessary to distinguish between a form word and a subject word when indexing for ERIC?

The form word may be part of a request specification.	8
If the distinction is not made, irrelevant material may be retrieved.	10
For both these reasons.	12

(from page 2)

Good. Now we can begin to look at concept indexing in detail.

Concept Indexing

The language we shall be dealing with is the language of educational research, and the selection of candidate terms should consist of those terms from the document which best convey its subject content.

We must remember that some of the terms and expressions we find in a particular document will be changed in translating into ERIC terms. But for the sake of exhaustivity, and in order to avoid interruption of thought and attention, it is best that we first write down all the concepts that we think are significant.

The best way to do this is by examining one section of the document at a time. For our demonstration documents in this course we have selected mostly short documents, but you must keep in mind that in indexing for ERIC you will be dealing with documents of varying lengths and complexity. The shortest articles may be dealt with paragraph by paragraph. In longer documents, the headings of the text may provide useful divisions. In most cases a dozen divisions should be enough.

Selection of candidate terms is best done by which of the following?

Using the terms most likely to be found in ERICT.	11
Using the language of the document.	9
Using ERICT as you go along.	3

(from page 6)

You say it is necessary to distinguish between a form word and a subject word because the form word may be part of a request specification.

This is part of the answer. For example, the request may be "I want to see a *bibliography* on programmed instruction."

But what happens if the distinction is not made? Return to 6 and select the *best* answer.

(from pages 11, 3, 7)

You feel that the selection of candidate terms is best done by using the language of the document, not ERICT. Good.

Since indexing for ERIC must be done in the least time consistent with high quality indexing, the best procedure is to select the significant concepts (both implicit and explicit) as you read through the document. You may use the same terms as the document since at this stage you are not concerned with whether they will be accepted in the original form or altered by ERICT. This method is quick and does not interrupt your thoughts and attention.

Depending upon the length of the document, you may have between 10 and 30 terms.

Before we consider how to recognize significant concepts, it should be noted that examination of the document is usually a mixture of concentrated reading and of relatively fast scanning. Our objective is good quality indexing at a minimum cost of the indexer's time.

How should a document be read for indexing purposes?

By first scanning it rapidly, then reading it thoroughly.	4
The balance of reading and rapid scanning will vary according to the content of the document.	6

(from page 6)

You say it is necessary to distinguish between a form word and a subject word because if the distinction is not made, irrelevant material may be retrieved.

This is part of the answer. If the request for a bibliography of programmed instruction resulted in a search for the two terms, Bibliography x Programmed Instruction, it could produce irrelevant material on Programmed Instruction for Bibliography. This would only be a problem, of course, if the form word was part of the request specification, so return to 6 and select the best answer to this question.

(from pages 7, 3)

No. This practice is dangerous since it may restrict your thinking. If you attempt to anticipate ERICT you may miss significant concepts in the document which you cannot instinctively think of in terms of ERICT.

It is important first to recognize the significant concepts in the document's own language. Later these can be translated into the language of ERIC.

Would you select useful candidate terms by using ERICT as you go along?

Yes	3
No	9

(from page 6)

Good. It is necessary to distinguish between a form word and subject word when indexing for ERIC because the form word may be part of a request specification, and if the distinction is not made, irrelevant material may be retrieved. For example, if the request is for a bibliography on programmed instruction, and the distinction between form words and subject words was not made during indexing, a search for the two terms "Bibliography" x "Programmed instruction" could produce an article entitled "Programmed Instruction for Bibliography", which would be irrelevant.

The major form categories for ERIC documents are:

(1) Descriptions of controlled experiments.

(2) Surveys (questionnaire, statistical, etc.)

(3) Case studies

(4) Reviews or summaries of research.

Definitions of these categories are to be found in Appendix B.

There are a few other forms, such as Programmed Text, that will occur far less frequently. Remember, in addition, that there will be many documents that need no form word.

Although the ERICT includes words that could be form words, they should not be used for this purpose. The words in the thesaurus are subject words. Specifying the form of a document is really part of descriptive cataloging. Although this course is concerned only with subject indexing, we refer to form specification since it is easy to confuse the two things.

Is it correct to say that the indexer should first of all allocate the document to one of the four categories listed in this lesson?

Yes	14
No	15

Your answer is that words which distinguish classes of documents are keywords. This is correct.

The Center for Documentation and Communication Research at Western Reserve University has developed a set of rules for indexing the literature of educational research. A prominent feature of these rules is a categorization of the main kinds of concepts to be found in the subject. The categories are:

a. Persons being studied. The terms here may be divided in various ways, e.g. by educational role (student, teacher, administrator, etc.); (adolescent, adult, etc.); by ability (gifted, slow learners, etc.).

b. Subject taught. (Physics, Psychology, etc.)

c. Major processes. (learning, achievement, adjustment, etc.)

d. Major operations. (teaching, grouping, testing, evaluating, etc.)

e. Teaching techniques. (lecture, programmed instruction, etc.)

f. Equipment and other agents of these techniques. (audio-visual aids, teaching machines, etc.)

g. Kinds, Parts, and Properties of these agents. (filmstrips, charts as kinds of audio-visual aids; sound track or color as part or property of filmstrips.)

The function of classification is to help us to see connections (relations) between things. Recognition of the category into which a term falls will help the indexer in a number of ways. For example, input unit, storage unit, and output unit are all parts of a computer. In indexing, it may be sufficient to use the term "computer" to stand for its parts.

If a document refers to blind, deaf and crippled children, but to none of them in great detail, which term best describes this concept?

Handicapped	17
Children	19
Physically Handicapped Children	21

(from page 12)

Not quite. The indexer should first of all see if the document falls into one of the four categories, but it may not belong to any of them. In any case the four categories listed here do not exhaust the form words that may be used now or in the future. But it is important to see first if the document belongs to a form category in order to distinguish a form word from the same term used as a subject word.

Now turn to 15 and go on with the lesson.

(from page 12)

Right. A document may not belong to any of the categories, but it is important to find out first so that you do not use subject words from ERICT to represent form categories.

As we mentioned before, the form words will be in a minority — the vast majority of terms will be subject words, to which we can now turn our attention.

As you are indexing, you should be trying to provide each document with keywords that will lead to its retrieval only when it is relevant to the question asked.

First, the keywords should be good "class words", i.e. they should be able to stand (alone or in combination with other words) for a class of documents which users of the collection find it useful to distinguish. For example, we have to index an article called "A Study of Certain Attributes of 45 High-School Varsity Football Team Members by Use of Psychological Test Scores."

You will probably dismiss (without even thinking about it) terms like "A", "Of", "Certain", "45", "By", and "Use". None of these constitutes a useful keyword, either by itself or with another word.

One way of determining whether a term can reasonably play the part of a class term is to ask yourself: Does the term fit into a phrase such as: "This document deals, to some extent, with the subject _X_." If a term can replace the "X", it is probably one you should consider for indexing. You can see at once that "X" could easily be replaced by terms like "attributes", "high school", "varsity team", "football team", "Psychological test", or "test scores."

Keywords are which of the following?

Words that distinguish classes of documents. 13

Words that are found in the title of the document. 16

(from page 15)

Your answer is that words that are found in the title of the document are keywords.

Not necessarily. Although some titles are rich in keywords, others may not be. Remember that the keywords must give a clue to the concepts dealt with in the document. You may find that some titles are misleading or confusing, and you must be sure that the content of the document confirms the terms used in the title before these terms are used as keywords.

Keywords are words representing significant concepts that form recognizable classes, and are not words like "and" or "of" or "with". Keywords can usually replace "X" in the statement "This document, to some extent, deals with the subject X."

Now turn to 13 and go on with the lesson.

(from page 13)

You feel that the document about blind, deaf, and crippled children is best described by the indexing term "Handicapped".

No. The single term "handicapped" might retrieve documents on handicapped adults or on emotionally handicapped children or adults. The single term "children" might retrieve documents on gifted children, emotionally-disturbed children or others that would not be relevant to a request for material on blind, deaf, and crippled children. The realization that all three of these are types of physical handicaps in children allows us to use the term "Physically Handicapped Children", to convey the general tenor of the information.

If a document concerns colleges of electrical engineering, mechanical engineering, mining engineering, which term would you use?

Colleges	18
Engineering colleges	21

(from pages 17, 19)

No. You are still making the mistake of choosing too general a term. "Colleges" would retrieve documents on colleges of art and theological colleges as well as colleges of various kinds of engineering.

Now turn back to 17, read the lesson carefully and try the question again.

(from page 13)

You feel that the document about blind, deaf, and crippled children is best described by the indexing term "Children".

No. The single term "children" might retrieve documents on gifted children, emotionally-disturbed children or others that would not be relevant to a request for material on blind, deaf and crippled children. The realization that all three of these are types of physical handicaps in children allows us to use the term "physically handicapped children" to convey the general tenor of the information.

If a document deals with colleges of electrical engineering, mechanical engineering and mining engineering, which term would you use to describe it?

Colleges	18
Engineering colleges	21

(from page 23)

You say that you would refer the searcher to Buros' Mental Measurements Year Book.

But chances are that this new information has not yet been entered in the year book. Return to 23 and try again.

(from pages 13, 17, 19)

Yes. The one term "Physically Handicapped Children", for example, would cover the three kinds of handicaps in children discussed in this document. And it would tell the searcher that this document would not be relevant to a request for information on gifted children, emotionally-disturbed children, or on physically handicapped adults.

If you are aware of the various categories to which terms belong you will more easily recognize the significance of variant word forms. For example, such words as "teaching" and "teachers" or "testing" and "tests" reflect quite clearly the different categories to which they belong. "Teaching" is an action (a form of behavior), whereas "teachers" are persons who perform the action.

Since variant word forms represent different aspects of the same thing, there is often no reason for preferring one form to another. For example, two articles entitled "Training of Teachers" and "Training for Teaching" may be identical in content. Where possible the choice should be based on the emphasis in the document. The real difficulty is that while one may theoretically distinguish between two word forms, in most documents the discussion of one inevitably involves the other. The best advice we can give is always to use the verb form, unless there is a strong reason for preferring the noun.

If a document deals with the hobbies of teachers which word would be the best one to use as an indexing term?

Teaching	24
Teachers	22

(from page 21)

You say that the best form to use is the noun "teachers."

Right. Our advice is to use the verb form "teaching" unless there is a strong reason for using the noun form. In this case there is a reason for using the noun — the document deals with the hobbies of teachers, rather than with teaching itself. It is the person rather than the action that is dealt with in this document, so we would use "teachers" rather than "teaching."

New Information. Our main objective is to index new knowledge which is the result of research, but research always draws heavily on past work, on established data, theories, and procedures. In every article you will index you will probably find references to past work. In indexing you have to distinguish between the new knowledge and data and procedures which are mentioned to describe the way in which the study was conducted and the existing evidence it drew on. These latter concepts are not in themselves the subject of the document, so you would not use them as indexing terms.

For established knowledge there are standard reference works, and you should be familiar with the scope of those covering the subject of Education. A question on the nature of a particular test, for example, may be answered by use of Buros' Mental Measurements Year Book. If a document merely refers to the use of a test, without description, analysis or evaluation, you would probably decide not to include the test in your indexing. It is conceivable, however, that someone may ask for examples of experiments in which the test had been used. This information may not be available elsewhere, and an index entry may be required.

Decisions of this kind do not depend on indexing theory. They depend on your experience and judgement.

Turn to page 23.

(from page 22)

If a document not only referred to the use of the test, but also provided new information for someone proposing to use it, what would you do?

Refer them to Buros' *Mental Measurements Year Book.*	20
Use the *name* of the test as an index term.	26
Use the word test as an index term.	28

(from page 21)

You say that the best term to use is "teaching".

No. Our advice was to use the verb form unless there was a strong reason for preferring the noun. Here there is a strong reason, because it is the person, rather than the action, that is dealt with in the document, so in this case we would use the noun form — "teachers".

Now you may continue on 22.

The terms "teaching" and "reading" are considered more specific in a general library than they would be in an ERIC clearing house because information on teaching and reading is only a small part of the total collection.

Yes. In a collection devoted to education, "teaching" and "reading" are very general terms. Similarly, terms which in the context of the whole ERIC system are relatively specific, may be very general in any one ERIC Clearing House. For example, in a clearing house specializing in the disadvantaged, all the documents would deal either directly or indirectly with the disadvantaged. But you must keep in mind that you will be indexing for the entire ERIC system, and the term "disadvantaged" is still necessary outside that particular Center.

EXHAUSTIVITY

Very exhaustive indexing assigns terms to concepts even when they play quite minor roles in the document's subject matter, and in extreme cases as many as a hundred or more terms might be assigned. In conventional library catalogs, half a dozen terms are usually enough for books.

Most documents in the ERIC system are likely to need between 3 and 30 terms, with an average of twelve.

The basic criterion for an indexing term is the significance of a concept as an element in the total information content of the document. It is fairly easy to decide the main theme of a document, but beyond the first half-dozen keywords there may be differences of opinion.

At which stage of indexing for ERIC should you be concerned with exhaustivity?

During the concept indexing stage.	35
During the translation stage.	32
During both the concept indexing and translation stages.	34

(from page 23)

You say you would use the name of the test as an index term for this document.

Yes. As we shall explain shortly, the name of the test would be be listed as an identifier, and you may need to use the descriptor TEST as well. Particular statistical methods and analytical techniques may be treated in the same way.

Location of Terms in Document

Their position in the document helps us to recognize the relative significance of different terms.

The title is usually rich in keywords, although you must be careful of fanciful or misleading titles.

The summary or abstract which often follows the title is also relatively rich in keywords, some of which merely duplicate or echo those in the title.

The introductory paragraph often refers to the existence of other inquiries into similar problems. While these may be valuable in acquainting you with the general theme of the document, they should be treated with great caution as a source of keywords. Most of them will refer to the information content of other documents, not the one you are indexing. However, insofar as the introductory remarks refer to the objectives of the document you are indexing, they may be a valuable source of less obvious candidate terms (implicit concepts).

Continued on page 27.

(from page 26)

The body of the text should be scanned to reinforce your awareness of the document's subject. It is often broken up into sections, such as "Methods", "Discussion", and "Conclusions".

Word Frequency and Generality

In any specialized field fundamental terms occur with great frequency. If selection were based solely on frequency, these terms would be assigned to nearly every document in educational research; e.g. most of the documents in an ERIC clearing house inevitably refer to such concepts as "Education", "Teaching", "Learning", "Student", "School", etc. Such terms are not effective as distinctive labels. They are so general that they are shared by far too many documents to act as efficient "screening devices". You will find these words in ERICT usually in conjunction with other terms (e.g. SCHOOL ADMINISTRATION, SCHOOL ARCHITECTURE, SCHOOL ATTITUDES, etc.).

In a general library, the terms "teaching" and "reading" would be considered moderately specific, whereas in the ERIC clearing houses these are considered as extremely broad terms.

Why is this so?

Because information on teaching and reading is only a small part of the total collection in a general library.	25
Because the system of indexing in a general library is different from that in an ERIC center.	29

(from page 23)

You say that you would use the word "test" as an index term.

This is too broad a term and wouldn't tell the searcher which test is referred to in this particular document.

Return to 23 and try again.

(from page 27)

Your answer is that the terms "teaching" and "reading" would be considered more specific in a general library than in an ERIC clearing house because the system of indexing is different.

This is not the reason. The generality of a concept depends upon the number of documents in the collection which refer to it. A general library covers the whole of knowledge and comparatively few documents refer to education. ERIC clearing house collections deal directly or indirectly with educational research, where terms like "teaching" and "reading" are encountered frequently and must therefore be considered general terms.

Now turn to 25 and go on with the lesson.

(from page 35)

You feel that you must be concerned with specificity during the concept indexing stage.

This is partly right. As you are seeking candidate terms during the concept indexing stage, you should be as specific as the document itself allows. But the final choice of terms will depend upon ERICT, and you should be careful at this stage to choose terms from the thesaurus which translate the candidate terms as specifically as possible.

Now go back to 35 and try the question again.

(from page 35)

You feel that you must be concerned with specificity during both the concept indexing stage and the translation stage.

Right. During the concept indexing stage you must be as specific as the document allows. In the translation stage ERICT will help maintain specificity, even though you may sometimes find that a more general term could be used to replace some of the specific terms you have listed. The problem of maintaining specificity when ERICT fails to provide the exact term is discussed in the following section on translation.

TRANSLATION OF CONCEPT INDEXING

As we explained earlier, retrieval of information depends on matching indexing terms with inquiry terms. Its success is measured by how nearly it approaches the ideal result — production of all documents relevant to the question, and none that is irrelevant.

Matching is made difficult by the existence of synonyms and other closely related words. For example, a document on the connection between "Anxiety and scholastic achievement in children" may be indexed by the terms Anxiety, Scholastic achievement, Children. A searcher may express the same ideas as Worrying, Success, School, Child. Since there is no match between the two sets of terms the document would not be retrieved.

The ideal result is difficult to achieve because of the complex relationships between subjects. For example, there may be documents on the connection between other emotional factors and scholastic achievement, which have a bearing on the problem of anxiety.

These difficulties explain the need for a controlled language in indexing.

Would you be able to retrieve a document indexed as "High school mathematics teaching" by searching for High school x Mathematics x Teachers?

Yes	36
No	37

(from page 25)

You say you should be concerned with exhaustivity during the translation stage.

No. By the time you get to the translation stage you should have a list of candidate terms for all the concepts you want to index.

Please return to 25 and select a better answer.

(from page 35)

You feel that you must be concerned with specificity during the translation stage.

This is partly right. The final choice of terms will depend upon the ERICT and you should choose terms from the thesaurus which translate candidate terms as specifically as possible. But during the concept indexing stage you must recognize the degree of specificity in the document itself.

Now turn back to 35 and try the question again.

(from page 25)

You say you should be concerned with exhaustivity during both the concept indexing stage and the translation stage.

No. By the time you get to the translation stage, you should have a list of candidate terms for all the concepts you want to index. You will then translate them into the language of ERIC.

Turn back to 25 and try the question again.

(from page 25)

You say you should be concerned with exhaustivity during the concept indexing stage.

Yes. It is during the first phase of indexing — the concept indexing stage — that you should look for all the concepts (explicit and implicit) dealt with in the document and compile a list of "candidate terms".

SPECIFICITY

It is important that the chosen terms should be as specific as the document allows. If the document refers to a particular species of anything, you should prefer the name of that species to the name of its genus. This is specificity. You should prefer "Audio-visual aids" to "Instructional aids", "Public speaking" to "Speaking", "Mental Retardation" to "Retardation", "Blind" to "Visually Handicapped", etc.

The importance of specificity is that it allows precision in retrieval. If you use the term "Visually Handicapped" instead of "Blind", you lose the ability to pinpoint material on the subject of the blind. It will always be mixed up with other, possibly irrelevant material on other kinds of visually handicapped persons (nearsighted, squint, etc.).

Don't confuse this with our earlier discussion on the use of one term to cover a number of related terms which were not dealt with to any great extent in the document.

We stated before that we are concerned with exhaustivity during the concept indexing stage. When must we be concerned with specificity?

During the concept indexing stage.	30
During the translation stage.	33
During both the concept indexing and the translation stage.	31

(from page 31)

No, you would not be able to retrieve "High school mathematics teaching" by looking for High school, Mathematics and Teachers; you would have to use Teaching instead of Teachers to get a match between the two sets of terms.

This is one reason for using an authority list like ERICT; it decides which of several terms you should always use, or it refers you to similar terms (with slightly different meanings) that may enable you to retrieve relevant documents. In this way you simplify the task of matching the document description and the search prescription.

Now turn to 37 and go on with the lesson.

Good, now we can go on.

The ERIC Thesaurus

After the indexer has written out the results of his concept indexing, or the questioner his provisional question, he must translate these terms into the language of ERICT. This means he must look up each term to see:

1. If ERICT uses it, in which case it is accepted *as it stands*.
2. If ERICT lists it but prefers a *synonymous* term, in which case the latter is accepted instead.
3. If ERICT uses a pre-coordination which corresponds to your concept indexing, in which case the *pre-coordinated* term is used.
4. If ERICT indicates *related terms* that assist definition and interpretation and/or suggest further terms to be added. In this case use the most appropriate term or terms.

The introduction to ERICT describes in detail how to use the thesaurus and this must be read. Here, we shall simply explain the essentials of its mode of operation.

ERICT was the result of indexing nearly 2,000 documents on the education of the disadvantaged. Though it will grow steadily in size, *it will never contain terms that have not already been used for indexing*. In this way it differs from many authority lists (such as the classification schemes used in libraries) which *attempt* to list all the concepts that *could be* required for indexing.

If you fail to find a term in ERICT what is the most likely reason?

The term has not yet been used for indexing.	40
A synonym for the term has been used.	38

(from page 37)

No. Usually there is a reference from expected synonyms. The main reason for terms not being listed in ERICT is their absence from documents already indexed. New terms are added to the thesaurus only as they become necessary for indexing documents in the system.

Please continue on 40.

No. This entry means that the term NEWS MEDIA is always to be Used For the term Press.

Please continue on 42.

(from page 37)

Yes. New terms are added to ERICT only as they become necessary for indexing documents in the System. Recommending new terms will be referred to in Lesson 3.

Synonyms in ERIC Thesaurus

The usual practice of ERICT is to include all synonyms for a concept and to make clear which term is to be used. For example, your concept indexing may include the term DISABLED. If you look for it in ERICT you will find:

Disabled
USE PHYSICALLY HANDICAPPED

Such an entry is known as a cross reference and is printed in small capitals.

If you look at the entry for Physically handicapped you will find:

PHYSICALLY HANDICAPPED
UF Disabled.

UF means Used for. The entry here helps to define the scope of the descriptor PHYSICALLY HANDICAPPED, as well as repeating the information that the term Disabled is not to be used.

If you found this entry in ERICT
NEWS MEDIA
UF Press

which term would you use in indexing?

News Media	42
Press	39

Not quite. It is true that ERICT would use the plural here, but no reference is made from the singular form.

Please continue on 44.

(from page 40)

Yes. Terms to be used for indexing are printed in large capitals. The symbol UF indicates rejected synonyms.

Word form control in ERIC Thesaurus

This is limited to the choice of nouns rather than adjectives (Anxiety, not Anxious), and of either the plural or the singular of a given word. The singular form is used for nouns of which you would ask How Much? the plural for those of which you would ask How Many? For example: How much TALENT? How many CHILDREN? No reference is made from the rejected form.

Occasionally, both singular and plural forms are used when different aspects of the subject are intended. For example, "Organization" implies a generalized study of the structure of groups, "Organizations" implies the groups themselves.

Other word forms are not controlled by ERICT. For example, it includes both Tests and Testing, Teachers and Teaching, Delinquents and Delinquency, Labor and Laborers. These relationships are usually among those indicated by the symbol RT (Related term).

If your concept indexing included the term School, what would you expect to find when you looked up School in the Thesaurus?

SCHOOL	45
Nothing	44
School Use Schools	41

(from page 44)

Yes. The inclusion of Indo-European Languages as a BT indicates that GREEK is meant to include the Modern as well as the Classical language.

Non Generic relations in the ERIC THESAURUS.

The last set of references under PHYSICALLY HANDICAPPED reads

RT Perceptually Handicapped
Physical Characteristics
Physical Development
Physical Handicaps
Physical Therapy
Rehabilitation Counseling

RT means Related term, and indicates terms bearing a relationship, other than generic, to Physically Handicapped. Perceptually Handicapped is another species of the BT Handicapped. Physical Characteristics are properties and Physical Development is a process — referring to any person. Physical Handicaps are properties of the Physically Handicapped. Physical Therapy and Rehabilitation Counseling are actions performed for the Physically Handicapped.

By checking the RT's the indexer may find a more suitable term for his purpose. The searcher is likely to make more use of RT's, as we shall explain shortly.

If your concept indexing includes Methods of Measurement and you find the following entry in the Thesaurus, which term will you use for indexing?

MEASUREMENT
RT Calculation
Instrumentation
Measurement Goals
Measurement Instruments
Measurement Techniques
Operation Research
Rating Scales
Scoring
Etc.

Measurement	49
Measurement Techniques 5	50
Measurement Instruments	48

(from page 42)

Yes. ERICT uses the plural form for nouns of which you can ask "How many?" No reference is made from the singular; the entry would be SCHOOLS.

Generic Relations In ERIC Thesaurus

The entry for PHYSICALLY HANDICAPPED also includes the following references:

NT Crippled children
Orthopedically Handicapped
BT Handicapped

NT means narrower term, and indicates that Crippled Children and Orthopedically Handicapped are kinds of Physically Handicapped.

Although NT is used mainly for this purpose in ERICT it sometimes shows other relationships. For example, Electricity is not a kind of Physics and Sanitation is not a kind of Public Health.

BT means broader term, and indicates that Handicapped is a genus of which Physically Handicapped is a species.

These relationships help to define the term listed and assist you in deciding if it is the correct translation of the concept you had in mind. You should always check the NT list to make sure that there is not a specific term that more accurately translates your concept.

Generic relations are mainly used in searching. This will be described shortly.

If your concept indexing included "Modern Greek" and you found this entry in ERICT

GREEK
UF Classical Greek
BT Classical Languages
Indo-European Languages

what would you?

Use the term GREEK	43
Not use the term GREEK	51

(from page 42)

No. The general rule is that you use the plural form for nouns of which you can ask the question "How many?"

Please continue on 44.

(from pages 50, 52)

Yes. Whenever a pre-coordinated form is listed in ERICT you should prefer it to the isolated terms.

Identifiers.

ERICT does not include names of particular schools, tests, etc. When such names occur as important items in a document, they should be indexed. They are known as identifiers to distinguish them from the descriptors in ERICT.

There is one difficulty created by this distinction. Each identifier will be related to one or more descriptors, though ERICT does not show such relationships. For example, the identifier Harvard University is related to the descriptor UNIVERSITIES. The indexer must decide in each case whether there is a related descriptor that should also be used.

An important function of ERICT is to decide which of a number of similar terms should always be used, or to refer the indexer or the searcher to similar terms already in use (perhaps with a slightly different meaning). This is only one of a number of kinds of control that may be applied to an indexing system to aid matching the indexer's description of a document and the searcher's description of his inquiry.

Controls may be applied by the indexer or the searcher, or both. They are of two kinds: those which tend to improve the recall of relevant documents and those which tend to improve the precision of documents retrieved.

Please continue with the lesson on 47.

(from page 46)

CONTROLS BY THE INDEXER

1. Synonyms and variant word forms.

These two kinds of related words were introduced in Lesson 1 and you have just seen how they are treated in ERICT. Together they form the most important method of achieving a high recall ratio.

You have selected "Home conditions" as a candidate term at the concept indexing stage. ERICT includes the following entries:

ENVIRONMENT
NT Family environment (etc.)

FAMILY (sociological unit)
RT Family environment (etc.)

FAMILY ENVIRONMENT
UF Home conditions (etc.)
BT Environment (etc.)

Home Conditions
Use Family environment

Which index term would you use as a translation?

Environment	53
Family	56
Family environment	55
Home conditions	54

(from page 43)

Not quite. Measurement Instruments is certainly more precise than Measurement, but another term in the RT list, Measurement Techniques, is a more accurate translation of your original phrase Methods of Measurement.

Please continue on 50.

(from page 43)

No. The term Measurement Techniques is close to your Methods of Measurement. You should choose it rather than the broader term Measurement.

Please continue on 50.

(from page 43)

Yes. RT Measurement Techniques as a translation of your original phrase, Methods of Measurement, is the most accurate.

PRE-COORDINATION in ERIC Thesaurus

Sometimes terms appearing separately in ERICT may also appear as a pre-coordinated, composite term, e.g. Teachers and Attitudes appear together as Teacher Attitudes. The advantage of this practice is that it narrows the search and eliminates documents that include the two concepts associated in other, possibly irrelevant ways e.g. Attitudes towards Teachers, or even Attitudes towards something else in a context that also refers to Teachers. The indexer and the searcher can both benefit from this practice of pre-coordination, which is common in ERICT.

A document deals with the opinions of students. Which of the following terms would you use for indexing?

Student Opinion	46
Students and Opinions	52

No. The listing of Indo-European Languages as well as Classical Languages under BT indicates that the term GREEK includes Modern as well as Classical.

Please continue on 43.

(from page 50)

No. It is true that ERICT includes the terms Students and Opinions, but whenever a pre-coordinated form is listed you should always prefer it to the isolated terms.

Please continue on 46.

(from page 47)

No. You have chosen too general a term. If you had looked at the example carefully, you would have seen that it gave a reference to the narrower, more correct term.

Now go back to 47 and try again.

(from page 47)

No. You have used the same term that you had in mind at the concept indexing stage. This term is not used by ERICT as an indexing word, and it has an instruction telling you what the preferred synonym is.

Now go back to 47 and try the question again.

(from page 47)

Good, you are right.

2. Weighting.

Weighting is a device to indicate the relative importance of concepts indexed. The ERIC system uses a very simple form of weighting in which some terms are described as major, so that we can limit the documents that we retrieve to those in which a concept is a major theme.

For example, suppose the search was for documents whose major terms were writing, objectives, and programming (marked by an asterisk in the example). Of two documents:

1. Writing * / Objectives * ? Programming*
2. Writing / Objectives * / Programming *

document number one would be retrieved, since all three terms are major; number two would not be retrieved in this search because writing is not a major term. Weighting is therefore a device for improving the precision of retrieval.

Which of the following documents would be retrieved by a search for "Student achievement in language studies"? (major terms only)

1. Student achievement * / language studies
2. Student achievement * / language studies *
3. Student achievement / language studies *

1; 2; 3 59

2 57

(from page 47)

No. This is much too general a term. If you had studied the example you would have noticed that a reference pointed to the narrower, more correct term.

Now go back to 47 and try again.

Good. Now let us leave controls imposed by the indexer and look at controls introduced by the searcher.

CONTROL BY THE SEARCHER

However strict the control over index terms and their use exercised by ERICT and the indexer, there will be occasions on which the searcher has to introduce controls of his own. First we will deal with controls that broaden the search.

Synonym And Word Form Control

The fact that ERICT is growing all the time means that it cannot always be expected to control synonyms. There may be occasions on which a term used by the searcher will not reveal all the literature on that topic because some documents have been indexed under another term whose relationship with the first terms has so far gone unnoticed. Furthermore, the intentional use in ERICT of apparent synonyms, such as Teaching and Instruction, and of variant word forms, such as Delinquents and Delinquency, will have the effect of separating closely related material.

(Of course the indexer may introduce the control himself by always using only one of the terms in question, but unless he announces his decision in some way the searcher may legitimately choose one of the unused terms and retrieve nothing at all!)

The searcher, however, can restore the situation. When his search retrieves insufficient material, he may broaden it to correct the intentional or unintentional reduction of material at any one of the synonyms or variant word forms by including others. He should include those other terms in his search as alternatives to the first, unsatisfactory term (a logical sum).

Continue on 58.

(from page 57)

For example if he searches unsatisfactorily for Teaching x Children, he may broaden his search and retrieve more documents by adding a synonym: (Teaching + Instruction) x Children. If he searches unsatisfactorily for Delinquents x Anxiety x Achievement, he may broaden the search and retrieve more documents by adding a variant word form: (Delinquents + Delinquency) x Anxiety x Achievement.

If a search for "Counseling x Universities" is unsatisfactory and you discover the entry in ERICT

COUNSELING
UF Counseling techniques
NT Educational counseling (etc.)
RT Counseling services
Counselors (etc.)

which of the following would broaden the search by addition of a variant word form?

(Educational Counseling + Counseling) x Universities	62
(Counseling + Counselors) x Universities	66
(Counseling services + Counseling) x Universities	64
(Counseling techniques + Counseling) x Universities	60

(from page 55)

No. You have ignored the asterisks (or the lack of them) that indicate weighting. You should have chosen those documents whose descriptions included the terms in question only if those terms had an asterisk.

Now go on to 57.

(from page 58)

No. You have chosen to add a synonym that ERICT has decided not to use. You are not introducing a control by doing this; you are overriding a useful control already imposed by ERICT. You should have added a variant word form.

Please return to 58, read the lesson carefully and select another answer.

(from pages 63, 66)

No. Although you have certainly broadened the search by using a higher generic term, it will retrieve only documents indexed by that term, and presumably dealing with the topic in general. We want to be able to retrieve not only documents of this general kind, but also documents dealing with all the various specific terms to supplement our knowledge of the particular kind we were originally interested in. We shall need to search not only for the genus, but also for the other species.

Now try again.

Which of the following best broadens an unsatisfactory search for Teaching machines x College libraries?

Teaching machines x Libraries	67
Teaching machines x (College libraries + Classroom Libraries + Public libraries + School libraries)	63
Teaching machines x (College Libraries + Classroom libraries + Public libraries + School libraries + Libraries)	65

(from page 58)

No. You have chosen to add a narrower term. You should have chosen a variant word form.

Please return to 58, read the lesson carefully and select another answer.

(from pages 61, 66)

You are nearly right. You have correctly decided to broaden the search by including other species but you should also add the genus.

Now try again.

Which of the following best broadens an unsatisfactory search for Objectives x Suburban Youth?

Objectives x (Suburban youth + Urban youth + rural youth + immigrant youth + Youth)	65
Objectives x Youth	61
Objectives x (Suburban youth + Urban youth + rural youth + immigrant youth)	67

(from page 58)

Not quite. You have chosen to add a related term, but it is not a variant word form.

Please return to 58, read the lesson carefully, and select another answer.

(from pages 61, 63, 66)

Yes. A generic search includes the genus and all its species.

NON-GENERIC SEARCH

Sometimes it is difficult to ask for *generically* broader terms, and instead we must use *non-generically* related terms, which may be properties, processes, agents of the unsatisfactory term. For example if a search for Anxiety x Academic Achievement does not retrieve enough material, we may alter the search to Anxiety x (Academic achievement + Gifted + High achievers) even though the added terms are not kinds of activity (as is Academic achievement).

If a search for Aptitude x Measurement goals proved unsatisfactory, how would you broaden it for a non-generic search?

Aptitude x Measurement	68
Aptitude x (Measurement goals + Measurement techniques + Testing)	70

(from page 58)

Yes. By using alternatives (in the form of a logical sum) the searcher can control synonyms and variant word forms.

GENERIC SEARCH

Sometimes the searcher may phrase his request quite properly and still retrieve insufficient material. He may then broaden the search by adding the genus and all its species. For example if he found that Anxiety x Junior High School x Achievement were unsatisfactory he could replace Anxiety by Psychological patterns or any kind of Psychological pattern (Anxiety, Fantasy, Fear, Insecurity).

If you felt that Junior High School was the inhibiting factor in the search we have just described, which of the following would you choose to conduct a generic search?

Anxiety x High School x Achievement	61
Anxiety x (Junior High School + Senior High School) x Achievement	63
Anxiety x (High School + Junior High School + Senior High School) x Achievement	65

No. You have just made the same mistake twice. The object of broadening the search in a generic hierarchy is to add to the search both the other species in the genus, and the name of the genus itself. In this way you will retrieve documents that supplement and include material on your original topic.

Now go back to 61 and read the lesson carefully. Then try the question again.

(from page 65)

No. You have altered the search simply to use the more general term that includes the original, unsatisfactory term. As we saw a moment ago, we should include specific terms closely related to the original term.

Now try again.

If a search for Student attitudes x Slum conditions proves to be unsatisfactory, how would you broaden it for a non-generic search?

Student attitudes x Slums	71
Student attitudes x (Slum conditions + Slum schools + Cultural disadvantagement)	70

(from page 70)

No. Instead of adding a term you have substituted Research for Mathematics. This is also likely to produce many irrelevant documents. To increase precision of retrieval the most effective method is adding more terms to the search prescription. In this case you could use Research x Mathematics x Universities.

Please continue with the Quiz on 73.

(from pages 65, 68)

Good, you are right.

CO-ORDINATION

The controls so far described (synonyms, word forms, generic and non-generic search) all imply looking for related words in ERICT. The effect in all cases is to broaden the search and thus increase the chance of retrieving documents relevant to the inquiry.

If we find that too many documents are retrieved, we need to contract our search and be more selective. It may be that the question was phrased rather loosely and can therefore be made more precise. But often it is simply an indication that the inquirer does not know how much literature there is on a subject.

It is only when an excessively large number of documents is retrieved that the inquirer realizes he can afford to be more precise in defining his need. The chief method for this is coordination i.e. demanding more terms in the indexing if documents are to be accepted.

Academic achievement may produce some 100 articles, and the searcher realizes that there is more literature on the subject than he was aware of. His real interest is in the problem only as it appears in High schools, so he reframes his search to include the additional term High schools: (Academic achievement x High schools). Now all of those documents on Academic achievement which do not refer specifically to High schools will be excluded. The search has been narrowed, irrelevant documents are rejected, and precision is improved.

You are interested in Research on Mathematics in Universities. A search for Mathematics x Universities produces too many documents. How can you control the search more effectively?

Research x Universities	69
Research x Mathematics x Universities	73

No. You have made the same mistake again. In a non-generic search the prescription is broadened by including related terms.

Now go back to 68 and read the lesson carefully. Then try the question again.

(from page 79)

YOUR ANSWER: Terms that occur frequently in documents are by themselves effective indexing terms.

No. Go back to 27 and review the section headed "Word frequency and generality". Then continue with the quiz on 82.

(from page 69)

Yes. The addition of extra terms to the search prescription is the most effective method of increasing precision.

Please continue with the quiz.

QUIZ

Question 1.

Which of the following statements is correct?

Concept indexing is independent of the index language.	75
Concept indexing is dependent on the index language.	77

(from page 80)

YOUR ANSWER: Persons studied, Subjects taught, and Educational processes are all examples of keywords.

No. You should review the descriptions of Keywords and categories on 15 and 13. Then continue with the quiz on 79.

(from page 73)

YOUR ANSWER: Concept indexing is independent of the index language.

Correct. Concept indexing should be carried out without reference to the index language. Please continue with the next question.

Question 2.

Surveys, Case studies, Experiments, Research are

Form words	76
Subject words	78
Either form words or Subject words	80

(from page 75)

YOUR ANSWER: Surveys, Case studies, Experiments and Research are form words.

This is partly right, but you should read the explanation again. Go back to 6 and 12, then return to 75 and try the question again.

(from page 73)

YOUR ANSWER. Concept indexing is dependent on the index language.

No. You need some revision. Go back to 7 and 9 and read the lesson again. Then go on with the quiz on 75.

(from page 75)

YOUR ANSWER: Surveys, Case studies, Experiments and Research are subject words.

This is partly right, but you should read the explanation again. Go back to 6 and 12, then return to 75 and try the question again.

(from page 80)

YOUR ANSWER: Persons studies, Subject taught, and Educational processes are all examples of categories of terms.

Correct. Please continue with the next question.

Question 4.

Terms that occur frequently in documents are by themselves

Effective indexing terms	72
Ineffective indexing terms	82

(from page 75)

YOUR ANSWER: Surveys, Case studies, Experiments and Research may be either form words or subject words.

Correct.

Please continue with the next question.

Question 3.

Persons studied, Subjects taught, and Educational processes are all examples of

Keywords	74
Categories of terms	79

(from page 83)

YOUR ANSWER: Exhaustivity is determined during concept indexing; Specificity during concept indexing and translation.

Correct. Please continue with the next question.

Question 6.

Which of the following is true?

ERIC Thesaurus exercises complete control over synonyms and variant word forms	86
ERIC Thesaurus exercises some control over synonyms and variant word forms	84

(from page 79)

YOUR ANSWER: Terms that occur frequently in documents are not by themselves effective indexing terms.

Correct. Please continue with the next question.

Question 5.

Which of the following is true?

Exhaustivity is determined during concept indexing; specificity during concept indexing and translation.	81
Exhaustivity is determined during translation; Specificity during concept indexing.	83

(from page 82)

YOUR ANSWER: Exhaustivity is determined during translation; Specificity during concept indexing.

No. You need to review the discussions of exhaustivity and specificity on 25 and 35. Then continue with the quiz on 81.

(from page 81)

YOUR ANSWER: ERIC Thesaurus exercises some control over synonyms and variant word forms.

Correct. Please continue with the next question.

Question 7.

Which of the following is true?

ERIC Thesaurus contains all the terms likely to be needed for indexing educational research documents.	87
ERIC Thesaurus contains only terms that have been used for indexing documents in the system.	88

(from page 88)

YOUR ANSWER: The symbol NT in ERIC Thesaurus always indicates species.

No. It sometimes shows other relationships. For example, Electricity is not a kind of Physics, and Sanitation is not a kind of Public Health.

Please continue with the quiz on 89.

(from page 81)

YOUR ANSWER: ERIC Thesaurus exercises complete control over synonyms and variant word forms.

No. You should review 40, 42 and 57. Then continue with the quiz on 84.

(from page 84)

YOUR ANSWER: ERIC Thesaurus contains all the terms likely to be needed for indexing educational research documents.

No. ERIC Thesaurus contains only terms that have already been used for indexing documents.

Please continue with the quiz on 88.

(from page 84)

YOUR ANSWER: ERIC Thesaurus contains only terms that have been used for indexing documents in the system.

Correct. Please continue with the next question.

Question 8.

Which of the following is true?

The symbol NT in ERIC Thesaurus usually indicates species.	89
The symbol NT in ERIC Thesaurus always indicates species.	85

(from page 88)

YOUR ANSWER: The symbol NT in ERIC Thesaurus usually indicates species.

Correct. Please continue with the next question.

Question 9.

Identifiers

Are included in ERIC Thesaurus	92
Are not included in ERIC Thesaurus	91

(from page 91)

YOUR ANSWER: Weighting is a recall device.

No. Go back to 55 and review the lesson. Then continue with Lesson 3.

YOUR ANSWER: Identifiers are not included in the ERIC Thesaurus.

Correct. There is one more question.

Question 10.

Weighting is a

Precision device	93
Recall device	90

(from page 89)

YOUR ANSWER: Identifiers are included in ERIC Thesaurus.

No. Proper names (of tests, schools, etc.) are known as identifiers. They are not listed in ERICT.

Please continue with the quiz on 91.

YOUR ANSWER: Weighting is a precision device.

Correct.

You may now go on to Lesson 3.